A SEMIGROUP APPROACH TO THE JUSTIFICATION OF KINETIC THEORY

KARSTEN MATTHIES AND FLORIAN THEIL

Abstract. This paper develops a method to rigorously show the validity of continuum description for the deterministic dynamics of many interacting particles with random initial data. We consider a hard sphere flow where particles are removed after the first collision. A fixed number of particles is drawn randomly according to an initial density $f_0(u, v)$ depending on d-dimensional position u and velocity v. In the Boltzmann Grad scaling, we derive the validity of a Boltzmann equation without gain term for arbitrary long times, when we assume finiteness of moments up to order two and initial data that are L^∞ in space. We characterize the many particle flow by collision trees which encode possible collisions. The convergence of the many-particle dynamics to the Boltzmann dynamics is achieved via the convergence of associated probability measures on collision trees. These probability measures satisfy nonlinear Kolmogorov equations, which are shown to be well-posed by semigroup methods.

1. Introduction

Deriving continuum models as a scaling limit of atomistic particle dynamics is a fundamental problem of mathematical physics. The aim is to prove the validity of continuum equations like the Boltzmann equation to describe the effective behavior of many particle dynamics. The first rigorous derivation was given by Lanford [Lan75] for short times using the BBGKY hierarchy. The problem of convergence of this hierarchy was partially overcome by using sufficiently small initial data on unbounded domains [IP89] or by considering linear variants, related results can be found in [Spo78, BBS83, Spo91, CIP94].

In this paper we consider kinetic annihilation, a simplification of hard ball dynamics which keeps two central features of the original evolution: The initial state is random, the evolution is deterministic. We assume that the initial configuration of n particles in the phase space $U \times \mathbb{R}^d$ ($U = \mathbb{T}^d$ is the unit torus) are drawn independently with some density $f_0 \in L^1(U \times \mathbb{R}^d)$. As long as they are intact the centres of the spheres move along straight lines with constant velocity. When the centres of two spheres, which are still intact, come within distance a, then both spheres are destroyed. Kinetic (or ballistic) annihilation has been studied extensively in the physics literature, see [CDPTW03, Pia95, DFPR95, PTD02], including a proof that the Boltzmann approximation does not hold in one space dimension [EF85]. The model can be used e.g. to model growth and coarsening of surfaces, see [KS88]. The BBGKY hierarchy could be used to prove the validity of the continuum description, but it would require stronger assumptions on the initial distributions.

In this article we develop a novel approach based on semigroup theory. A preliminary version was first introduced in [MT10]. We consider heterogenous initial distributions and have to deal with the transport term in the emerging equation. Furthermore despite dropping assumptions on Poissonian distributions, we are able to show stronger convergence results.

We analyze the evolution of n balls of diameter a and with position $u(i, t) \in U \subset \mathbb{R}^d$ for $i \in \{1, \ldots, n\}$ with $d \geq 2$ and respective velocity $v(i, t) \in \mathbb{R}^d$. Our main interest is the kinetic limit, when the number of particles n tends to infinity and the initial values $(u(i, 0), v(i, 0))$ are independent identically distributed random variables distributed according to some initial distribution f_0. The diameter a of the particles is coupled to the number n by the Boltzmann-Grad scaling, which is in the easiest form

$$na^{d-1} = 1. \tag{1.1}$$

The final aim is to analyze the situation, where the particles interact via some suitable potential, like a hard-sphere one. Then for all open sets $A \subset U \times \mathbb{R}^d$ and any given time, the number of particles in A divided by the total number of particles converges to $\int_A f(u,v,t)\,\mathrm{d}u\,\mathrm{d}v$. It is expected that the density f then solves the nonlinear Boltzmann equation

$$\partial_t f + v \cdot \nabla_u f = -Q_-[f,f], \tag{1.2}$$

where $Q_- \geq 0$ is the collision operator accounting for the losses. The collision operator can be easily derived for hard-core potentials in a situation of completely independent particles with density $f(u,v,t)$. Particles with velocities v and v' collide at position u with a given probability depending on v and v' and impact parameter $\nu \in S^{d-1}$. In the density there is a loss at (u,v) and (u,v'). The loss-operator has the form

$$Q_-[f,f](u,v) = \int_{\mathbb{R}^d} \int_{S^{d-1}} f(u,v')\, f(u,v)\, [(v-v')\cdot \nu]_+ \,\mathrm{d}\nu\,\mathrm{d}v'.$$

In the case of collisional dynamics the loss is balanced by the corresponding gain at (u, v_*) and (u, v'_*) with the consequence that the Boltzmann equation is augmented by the appropriate gain-term.

In [MT10] we analyze a situation of spatially homogeneous initial data, which corresponds to a version of (1.2), where the transport term $v \cdot \nabla_u f$ vanishes. In the present paper we will allow for heterogeneous initial data reintroducing the transport term. To handle the transport term we will restrict our attention to initial densities f_0 which are absolutely continuous with respect to the Lebesgue measure:

$$f_0 \in L^1(U \times \mathbb{R}^d), \text{ with } f_0 \geq 0 \text{ and } \int_{U \times \mathbb{R}^d} f_0(u,v)\,\mathrm{d}u\,\mathrm{d}v = 1 \tag{1.3}$$

and have finite total energy:

$$\int_{U \times \mathbb{R}^d} (1+|v|)^2 f_0(u,v)\,\mathrm{d}u\,\mathrm{d}v = K < \infty. \tag{1.4}$$

We require the u-marginal of f_0 to be in $L^d(U)$ to ensure that particles overlap at any given point only with probability zero, see Section 7 and that the energy density and its transported versions are also bounded in $L^\infty(U)$

$$K_\infty = \text{ess sup}_{(t,u)\in(0,T)\times U} \int_{\mathbb{R}^d} (1+|v|^2)\, f_0(u-tv,v)\,\mathrm{d}v < \infty. \tag{1.5}$$

The main result (Theorem 2.1) is a rigorous justification of (1.2), if f_0 fulfills (1.3), (1.4) and (1.5). A key element in the proof is an intermediate layer of description between the complicated n-body evolution and the one-body distribution $f(\,\cdot\,,\,\cdot\,,t)$. This layer consists of trees which describe the history of collisions of an individual particle and its potential scattering particles. This extra layer allows on one hand a relatively easy description of the limiting (idealized) distribution P_t, see the definition in (4.34). On the other hand we can estimate the error between the empiric distribution $\hat{P}_t$ created by the n-body evolution and the idealized distribution P_t, see Proposition 5.4.

In contrast to [MT10], where we used explicit formulae for the distributions, we derive nonlinear Kolmogorov equations for the evolution of the probability measures P_t and $\hat{P}_t$ with time t. As we are essentially describing the evolution of low-dimensional marginals the Kolmogorov equations are quadratic in the measure. A key result is the derivation of the Kolmogorov equation for $\hat{P}_t$ which accounts for the correlations caused by the history of the evolution.

By fixing one of the arguments of the Kolmogorov operator it will be possible to apply general semigroup theory to the idealized evolution. A fixed-point argument then provides the existince of a nonlinear semigroup. The desired convergence of the multi-body empiric distribution to the idealized one in the Boltzmann-Grad limit then follows with relative ease. The final step is to derive the density description $f_t(\,\cdot\,,\,\cdot\,)$ as a marginal from the distribution of trees. The Boltzmann equation will appear then naturally from the differential equation for the distribution on trees.

Allowing heterogenous initial data requires a number of additions to the methods in [MT10], because several new error terms are created by the spatial heterogeneity. A careful analysis of regularity of the initial data (1.3) f_0 is needed to deal with concentration phenomena in position space and to obtain solutions to (1.2). We will consider here a bounded domain with periodic boundary data, i.e. $U = \mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$.

On the level of partial differential equations several formulations of (1.2) are relevant. In a particular form this equation is well-defined for L^1-data with respect to the space coordinate u. We require higher spatial regularity in the derivation, such that we can obtain standard mild and weak solutions of (1.2). Following ideas in [MT10], L^2 regularity is enough to prove tightness of the self-similar tree measure by deriving bounds on the expected number of nodes in the trees. We need L^d to obtain good bounds on the initial overlap of particles. In the current paper we impose L^∞ assumptions on the spatial energy density for simplicity of presentation.

It is noteworthy that the well-posedness of the Boltzmann equation in some function space does not imply that the limit of the single-particle distribution is a solution of the Boltzmann equation; an explicit counterexample has been constructed in [MT10].

In the current paper we prove all required regularity for finite times through a simple a priori bound of the solutions of (1.2) due to the sign of the right-hand side. However, we expect that with growing complexity more involved estimates will be required. The analytical understanding of various aspects of kinetic equations has progressed significantly within the last 25 years. A crucial tool is the gain of regularity and compactness through velocity averaging lemmas for various equations [Ag84, GLPS99, Ge90, GG92, DP01], for further references see the review [Per04]. The existence of renormalized solutions to the full Boltzmann equation [DL89a] uses transport theory as in [DL89b]. These tools are also relevant when the aim is to derive the incompressible Navier Stokes equation through scaling of solutions of the Boltzmann equation [BGL93, LM01a, LM01b, GS04].

The paper is organized as follows. In Section 2 we will describe the set-up and formulate the main result. In Section 3 the collision trees and various probability distributions are intoduced. The main theorem will be proved in Section 6, by deriving the effective single particle dynamics. In Section 7 we discuss spatial concentrations.

2. Set-up and main result

We define the multi-body evolution in the following way. Let $f_0 \in L^1(U \times \mathbb{R}^d)$ be a density of initial conditions. For each $n \in \mathbb{N}$ consider the random variable

$$(u^0, v^0) = (z_1, \dots, z_n) \in (U \times \mathbb{R}^d)^n, \tag{2.1}$$

with $z_1, \dots, z_n$ independent, identically distributed according to f_0 and a determined by (2.8) giving a probability measure Prob_a. The particles evolve by force-free Newtonian dynamics with initial conditions $\left(u_i^0, v_i^0\right) \in U \times \mathbb{R}^d$ for $i = 1, \dots, n$

$$u_i(t=0) = u_i^0, \qquad v_i(t=0) = v_i^0 \tag{2.2}$$

according to the differential equations

$$\begin{aligned} \dot{u}_i(t) &= v_i(t) \\ \dot{v}_i(t) &= 0 \end{aligned} \tag{2.3}$$

The scattering state (1 for unscattered, 0 for scattered and removed) for each particle $i = 1, \dots, n$ and time t is defined by

$$\beta_i^{(a)}(t) = \begin{cases} 0 & \text{if there exists } i' \neq i \text{ such that } |u_i^0 - u_{i'}^0| \leq a, \\ 0 & \text{if } \min\left\{ d(z_i, z_{i'}, s) - a\beta_{i'}^{(a)}(s) \;\middle|\; s \in [0,t], i \neq i' \right\} \leq 0, \\ 1 & \text{else,} \end{cases} \tag{2.4}$$

where a will depend on n and where the distance of particles on the torus U with data $z_i = (u_i, v_i)$ and $z_{i'} = (u_{i'}, v_{i'})$ is

$$d(z_i, z_{i'}, s) = |u_i(s) - u_{i'}(s)|_U = \left|u_i^0 - u_{i'}^0 + s\left(v_i^0 - v_{i'}^0\right)\right|_U .$$

This means in particular, that particles are removed if they overlap at time $t = 0$. See [MT10] for a proof, that $\beta_i^{(a)}(t)$ is well-defined. We compare the multi-body evolution with the single-body description $f : U \times \mathbb{R}^d \times [0, \infty) \to \mathbb{R}$

$$\begin{aligned} \partial_t f + v \cdot \nabla_u f &= -Q_-[f, f], \\ f(u, v, 0) &= f_0(u, v), \end{aligned} \tag{2.5}$$

where

$$Q_-[f, g](v) = L[g](v)\, f(v) = \left(\int_{\mathbb{R}^d} g(v')\, \kappa_d\, |v - v'|\, \mathrm{d}v' \right) f(v) \tag{2.6}$$

is the loss term and κ_d is the volume of the $(d-1)$ dimensional unit ball, in particular $\kappa_2 = 2$ and $\kappa_3 = \pi$. We will consider mild solutions of (2.5), which are functions $f \in C^0([0, T), L^1(U \times \mathbb{R}^d))$ with

$$f_t = S_t f_0 - \int_0^t S_{t-s} Q_-[f_s, f_s]\, \mathrm{d}s \tag{2.7}$$

in $L^1(U \times \mathbb{R}^d)$ for all $t \in [0, T)$ where S_t is the strongly continuous linear semigroup given by $S_t h(u, v) = h(u - tv, v)$.

Theorem 2.1. *Let $f_0 \in L^1(U \times \mathbb{R}^d)$ with $d \geq 2$ be an initial distribution fulfilling (1.3), (1.4), (1.5). For $n \in \mathbb{N}$, consider the evolution of (2.3) with initial conditions (2.2) as in (2.1). The diameter a is coupled to n via the Boltzmann-Grad scaling*

$$n a^{d-1} = 1. \tag{2.8}$$

Then the density of the unscattered particles converges to a solution of the Boltzmann equation in the sense that for all $\varepsilon > 0$ and all open $A \subset U \times \mathbb{R}^d$ uniformly for t in a compact set

$$\lim_{a \to 0} \mathrm{Prob}_a \left(\left| \frac{1}{n} \# \left\{ i \;\middle|\; (u_i(t), v_i(t)) \in A,\, \beta_i^{(a)}(t) = 1 \right\} - \int_A \mathrm{d} f_t(u, v) \right| > \varepsilon \right) = 0, \tag{2.9}$$

where $f_t(\,\cdot\,, \,\cdot\,) = f(\,\cdot\,, \,\cdot\,, t)$ is the unique mild solution of (2.5). Furthermore, there exists a sequence $a_k \to 0$ and corresponding particle numbers n_k, such that with probability 1

$$\frac{1}{n_k} \sum_{i=1}^{n} \beta_i^{(a_k)}(t)\, \delta(\,\cdot\, - (u_i(t), v_i(t))) \overset{*}{\rightharpoonup} f_t \tag{2.10}$$

weak- in $M(U \times \mathbb{R}^d)$ (the space of unsigned Radon measures) as $k \to \infty$, with δ denoting the Dirac-distribution.*

Remark 2.2.
- (i) *The number of particles n is fixed for given diameter a in difference to* [MT10]*, where it was a random number given by a Poisson distribution with intensity a^{1-d}. So here we consider a canonical ensemble as opposed to a grand canonical ensemble in the easier case. As we need some control of correlations to prove convergence in probability (2.9), proofs would not be much easier for a grand canonical ensemble.*
- (ii) *For other notions of solutions of equation (2.5) see Proposition 4.9.*
- (iii) *Some effects of spatial concentration are analyzed in section 7, concentration effects in velocity are ruled out via the absolute continuity with respect to Lebesgue measure.*
- (iv) *A larger class of initial distributions like $f_0 \in L^d(U, BC^0(\mathbb{R}^d)^*)$ with some additional non-concentration assumptions in velocity space seems to be conceivable, but is not considered for presentational reasons.*
- (v) *The convergence of k-particle distribution functions to a product of f_t can be shown for every fixed k using the same method. This gives a connection to the classical derivation for short times using the BBGKY hierarchy, which was applied to the simpler problem of coagulation by Lang and Nguyen* [LN80]*. Here the spheres move along Brownian paths and two intact spheres annihilate each other if the distance between the centers drops below a. Although the series generated by the BBGKY hierarchy does not converge globally, a rigorous justification of the corresponding Boltzmann equation was obtained*

by restarting the procedure at small positive time. The BBGKY hierarchy could also be applied to the ballistic annihilation model, but would require bounds on exponential momenta.

3. Collision trees

We introduce now the important intermediate layer of collision trees to understand the multi-body dynamics. In the collision tree with root (u, v), we will collect information of collisions and potential collisions up to time t for particle 1 with initial data (u, v). Collisions happen in the gainless case, considered here, if $|u + sv - (u' + sv')|_U \leq a$ for some time $s \in [0, t]$ and some $(u', v') = (u_i^0, v_i^0)$ for some particle i. So on the second level of nodes we will note all particle, which intersect the path of the root particle up to time t. In the tree we collect the velocity v_i, collision time s_i and impact parameter ν_i of the particle i. Such an intersection will only lead to a collision in (2.3), if both particles were not scattered before, e.g. the particle at the root might not collide at the smallest potential collision time, because this particle was scattered earlier. Therefore to determine this, we will also collect all information about potential collisions of the particles i on level 2 up to the collision time s_i. Of course, the potential collisions with i cannot occur, if there are earlier collisions of particles i' on level 3 before the collision time s_i. We will iterate this procedure to collect all possible collision events. Due to the finiteness of the number of particles and number of possible collisions in finite time for given velocities, the trees have finite size. We will later show, that the size of the trees relevant in the description of (2.3) is uniformly bounded as $n \to \infty$ in our scaling. To compare the dynamics of several particles, we will consider "trees" with α roots, which is a forest in graph theory language. The number α is fixed, in particular the behavior with $\alpha = n$ will not be considered. We use the following notation.

Definition 3.1 (Marked Trees)**.** *The set of trees $\mathcal{T}$ is the set of acyclic directed graphs $m = (V, E)$ with a root node 'root'. We adopt the standard convention that two trees $(V, E), (V', E') \in \mathcal{T}$ are identical if there there exists a bijective map $\varphi : V \to V'$ which maps the root to the root and conserves the connectivity ($(\varphi(l), \varphi(l')) \in E'$ if and only if $(l, l') \in E$).*
Each tree induces a partial order on the set of nodes. For a node $l \in m \setminus \mathrm{root}$ the predecessor of l is denoted by $\bar{l}$. We say that $l <_p k$ if either $l = \bar{k}$ or $l <_p \bar{k}$.
Let $Y = \mathbb{R}^d \times [0, \infty) \times S^{d-1}$ be the set of child-markers and $Y^* = U \times \mathbb{R}^d$ be the set of root markers. The collision trees $\mathcal{MT}$ is the set of mappings from trees to $Y \cup Y^*$ such that the collision times respect the order of the vertices:

$$\mathcal{MT} = \{(m, (u, v), (v_l, s_l, \nu_l)_{l \in m \setminus \mathrm{root}}) \ : \ m \in \mathcal{T} \text{ and } s_l < s_k \text{ if } l < k\}.$$

Each marked tree $\Phi \in \mathcal{MT}$ induces a partial order '$<$' on the set of vertices:

$$l < k \text{ if there exists } l' \leq_p l, k' \leq_p k \text{ such that } \bar{l'} = \bar{k'} \text{ and } s_{l'} < s_{k'}.$$

The number of vertices in a marked tree Φ is $\#\Phi$. The distance between two trees Φ and Ψ is defined as

$$d(\Phi, \Psi) = \begin{cases} \min\left\{1, \max\limits_{l \in m(\Phi)} |\Phi_l - \Psi_l|_\infty\right\} & \text{if } m(\Phi) = m(\Psi), \\ 1 & \text{else.} \end{cases}$$

By $\tau(\Phi)$ we denote the final collision time

$$\tau(\Phi) = \max\left\{s_l \ : \ \bar{l} = \mathrm{root}\right\},$$

and

$$\mathcal{MT}_t = \{\Phi \in \mathcal{MT} \ : \ \tau(\Phi) = t\}$$

is the set of trees where the final collision takes place at t. For each node $l \in m$ the initial position $u_l \in \mathbb{T}^d$ is computed via the recursive formula

$$u_l = \begin{cases} u & \text{if } l = \mathrm{root}, \\ u_{\bar{l}} + s_l(v_{\bar{l}} - v_l) + a\nu_l & \text{if } \bar{l} \text{ is the ancestor of } l. \end{cases} \tag{3.1}$$

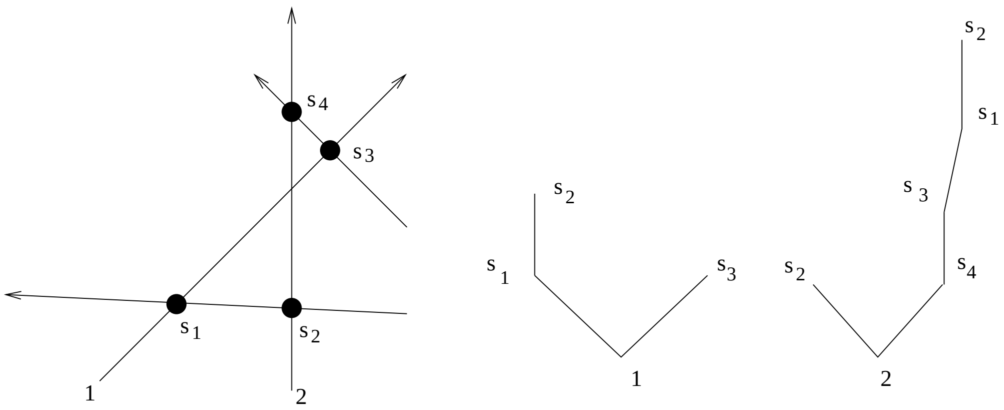

FIGURE 1. Initial positions and velocities of four particles. The bullets indicate the positions where the particles are potentially scattered. The collision at time s_4 can be ignored in the tree of with root particle 1, as $s_4 > s_3$ and $s_4 > s_2$. The tree with root particle 2 is atypical (or not 'good' in the sense below), because the collision event at time s_2 appears in separate branches of the tree.

We will often write $u(\Phi)$ and $v(\Phi)$ instead of $u_{\rm root}(\Phi)$ and $v_{\rm root}(\Phi)$ as well as τ instead of $\tau(\Phi)$.

Remark 3.2. *We can define a total order on the nodes for a generic subset of $\mathcal{MT}$. We consider trees $\Phi \in \mathcal{MT}$ if $s_l \neq s_k$ for all $l, k \in m(\Phi)$ such that $\bar{l} = \bar{k}$. It is easy to see that for each generic tree the relation '$<$' defines a total order and thus we can enumerate the nodes $l \in m(\Phi)$.*

We will also consider trees generated by several particles, in this case $\Phi \in \mathcal{MT}^\alpha$, $\alpha \in \{1, 2, \ldots\}$. The set $Y = \mathbb{R}^d \times [0, \infty) \times S^{d-1}$ denotes initial velocity v, collision time s and impact parameter ν. The root marker $Y^* = U \times \mathbb{R}^d$ characterizes the initial position and velocity of the root particle. Some examples of collision trees are given in figure 1. For a wider discussion of such tree structures see [MT08].
The definition of the evolution of the set of trees is based on two elementary operations: extraction of subtrees and pruning.

Definition 3.3. *Let $\Phi \in \mathcal{MT}_t$ (i.e. $\tau(\Phi) = t$) and let $l \in m$ be the node which corresponds to the final collision in the sense that $\bar{l} = \text{root}$ and $s_l = \tau$. The subtree $\Phi'_t = (m', (u', v'), (v'_k, s'_k, \nu'_k)_{k \in m'})$ is defined by*

$$m' = \{k \in m \ : \ k \leq_p l\}, \ (u', v') = (u_l, v_l), \ (v'_k, \nu'_k, s'_k) = (v_k, \nu_k, s_k) \ \text{if } k < l.$$

The pruned tree $\bar{\Phi}_t = (\bar{m}, (u, v), (v_k, \nu_k, s_k)_{k \in \bar{m}})$ is defined by $\bar{m} = m \setminus m'$ if $t = \tau$ and $\bar{\Phi}_t = \Phi$ if $t \neq \tau$.
The impact parameter ν_t is defined as $\nu_t := \nu_l = \frac{1}{a}(u - u' + t(v - v'))$ with the convention that $u = u(\Phi)$, $u' = u(\Phi'_t)$ etc.

Recall that $\mathcal{MT}$ is a metric space and denote for each $\Psi \in \mathcal{MT}$ by

$$B_h(\Psi) = \left\{\Phi \in \mathcal{MT} \ : \ d(\Phi, \Psi) \leq \tfrac{h}{2}\right\}.$$

the ball with diameter h centered at Ψ. For $0 < h < 1$ the ball $B_h(\Psi)$ is a $2d\#\Psi$-dimensional, smooth set.

Definition 3.4. *The standard Lebesgue-measure on $\mathcal{MT}$ is denoted by $\mathrm{d}\lambda$.*

We will now describe several probability measures on $\mathcal{MT}$ to first describe the idealized distribution P_t, closely related to the Boltzmann equation, and then the empiric distributions $\hat{P}_t$, related to the annihilation flow. We collect several properties of these to prepare Proposition 5.4 which delivers the convergence of $\hat{P}_t$ to P_t as $n \to \infty$.

4. Idealized distribution

The idealized distribution P_t is characterized by a differential equation (4.4). Before stating the equation we give a simple example which motivates the form and the analysis of the equation. Then we show that (4.4) admits a unique solution P_t. Finally we study the properties of P_t which will be instrumental when we demonstrate in section 5.1 that for each t the probability distribution P_t is very close to the empiric tree distribution $\hat{P}_t$, which is generated by the annihilation dynamics.
To motivate the analytical setting we consider first a simple example which illustrates the notation and the way semigroup theory applies. Recall that δ denotes the Dirac distribution and consider the linear system of differential equations

$$\left\{ \begin{array}{lll} \frac{\mathrm{d}}{\mathrm{d}t}u(t) & = \mu u, & u(0) = 1, \\ \frac{\partial}{\partial t}v(s,t) & = \delta(t-s)u + \mu v, & v(s,0) = 0 \text{ for all } s, \end{array} \right. \tag{4.1}$$

with a parameter $\mu \leq 0$ and time-dependent variables $(u, v) \in X = \mathbb{R} \times M([0,\infty))$. The Banach space $M(\Omega)$ is the set of all finite unsigned measures, or alternatively, the dual space of $C(\Omega)$. The solution is given by $u(t) = \exp(\mu t)$ and

$$v(s,t) = \left\{ \begin{array}{ll} \exp(\mu t) & \text{if } t \geq s, \\ 0 & \text{else.} \end{array} \right.$$

The generator takes the form $L_t = \begin{pmatrix} \mu & 0 \\ \delta(t- \cdot\,) & \mu \end{pmatrix}$, it is easy to see that L_t is stable for each $\mu < 0$, i.e. has a continuous resolvent for each $\lambda \geq 0$. Indeed, $(\lambda-\mu) f_1 = g_1$ and $-\delta(t- \cdot\,)\, f_1 + (\lambda-\mu) f_2 = g_2$ implies that

$$f_1 = \frac{1}{\lambda - \mu} g_1, \tag{4.2}$$

$$f_2 = \frac{1}{\lambda - \mu}\left(\frac{g_1}{\lambda - \mu}\, \delta(t - \,\cdot\,) - \mu g_2 \right), \tag{4.3}$$

which is clearly a continuous map from X to X. In the case of the example, the operator L_t is actually continuous

$$\begin{aligned} \|L_t f\|_X =& |\mu|\, |f_1| + \|f_1 \delta(t - \,\cdot\,) + \mu f_2\|_{M([0,\infty))} \\ \leq& |\mu|\, |f_1| + |f_1|\, \underbrace{\|\delta(t - \,\cdot\,)\|_{M([0,\infty))}}_{=1} + |\mu|\, \|f_2\|_{M([0,\infty))} \leq \max\{1, 2|\mu|\}\, \|f\|_X, \end{aligned}$$

this is not true for the operator which is defined below.
Now we consider a setting which is more closely linked with annihilation dynamics.

4.1. **Existence and uniqueness of the idealized distribution.** We now introduce the distribution P_t^a via the Kolmogorov equation

$$\left\{ \begin{array}{ll} \frac{\partial P_t^a}{\partial t} & = \mathcal{Q}_t^a[P_t^a, P_t^a, \mu_t^a[P_t^a]], \\ P_0^a & = f_0 \end{array} \right. \tag{4.4}$$

where the collision rate $\mu_t^a[P] \in M(\mathcal{MT})$ is defined by

$$\mu_t^a[P](\Phi) = \int_{S^{d-1}} \mathrm{d}\nu \int_{\mathcal{MT}} \mathrm{d}P(\Psi)\, \delta(u - u(\Psi) + t(v - v(\Psi)) + a\nu)\, [(v - v(\Psi)) \cdot \nu]_+, \tag{4.5}$$

and $\mathcal{Q}_t^a[P, P', \mu]$ is given by

$$\mathcal{Q}_t^a[P,\, P',\, \mu](\Phi) = P(\bar{\Phi}_t)\, L_t^a[P_t'](\Phi), \tag{4.6}$$

$$L_t^a[P',\, \mu](\Phi) = \delta(t - \tau(\Phi))\, P'(\Phi_t')\, [(v - v') \cdot \nu_t]_+ - \mu(\Phi), \tag{4.7}$$

with the convention $v = v(\Phi)$, $v' = v(\Phi_t')$ etc. Note that the operator L_t^a extracts subtrees which collide with the root particle. The initial position $u(\Phi_t')$ varies with a as in (3.1) and provides the sole mechanism how L^a depends on a. The dependency on a will be mostly suppressed in the subsequent analyis. As the dependency of $\mathcal{Q}$ on μ is rather trivial it will often not be shown

explicitly. The evolution defined by (4.6,4.7) will assign the probability $P(\bar{\Phi}_t)P'(\Phi'_t)\,[(v-v')\cdot\nu_t]_+$ to the new formed tree Φ at time t, while $\mu(\Phi)$ is the loss rate at later times.
We use the notation $P_t = P_t^0$ and refer to P_t as 'idealized' distribution.
For fixed Φ the time evolution of $P_t(\Phi)$ is rather simple:

$$P_t(\Phi) = 0 \text{ if } t < \tau(\Phi),$$

and for $t \geq \tau(\Phi)$ the function $t \mapsto P_t(\Phi)/P_\tau(\Phi)$ is non-negative and monotonously decreasing.
To see that $\mathcal{Q}_t$ conserves the total probability we have to show that the delta-distributions in (4.5) and (4.7) are equivalent.

Lemma 4.1. *Let* $\Phi \in \mathcal{MT}$ *then*

$$\delta(t-\tau(\Phi)) = \delta\left(u - u'(\Phi) + t(v - v'(\Phi)) + a\nu(\Phi)\right), \tag{4.8}$$

i.e.

$$\int_{\mathcal{MT}} \mathrm{d}\lambda(\Phi)\, g(\bar{\Phi}_\tau, \Phi'_\tau)\, \delta(t-\tau(\Phi))$$
$$= \int_{\mathcal{MT}} \mathrm{d}\lambda(\Phi) \int_{S^{d-1}} \mathrm{d}\nu \int_{\mathcal{MT}} \mathrm{d}\lambda(\tilde{\Phi})\, g(\Phi,\tilde{\Phi})\, \delta\left(u(\Phi) - u(\tilde{\Phi}) + t(v(\Phi) - v(\tilde{\Phi})) + a\nu\right).$$

for all $g \in C_c(\mathcal{MT}\times\mathcal{MT})$ *(the set of continuous functions with compact support) such that* $g = 0$ *if* $\max\{\tau,\tilde{\tau}\} > t$ *or* $(v-\tilde{v})\cdot(u-\tilde{u}+t(v-\tilde{v})) < 0$.

Proof. Let $\mathcal{MT}_0 = \{\Phi \in \mathcal{MT} \;:\; u_{\mathrm{root}} = 0\}$ and define for each $\Phi \in \mathcal{MT}$ and $u \in U$ the translated tree $\xi(\Phi,u) \in \mathcal{MT}$ as

$$\xi_l = \begin{cases} (u, v_{\mathrm{root}}) & \text{if } l = \text{root}, \\ \Phi_l & \text{else.} \end{cases}$$

Then we find that the left hand side of (4.8) can be rewritten as

$$\int_{\mathcal{MT}} \mathrm{d}\lambda(\Phi)\, g(\bar{\Phi}_\tau, \Phi'_\tau)\, \delta(\tau - t)$$
$$= \int_{\mathcal{MT}} \mathrm{d}\lambda(\Phi) \int_{\mathcal{MT}_0} \mathrm{d}\lambda(\tilde{\Phi}) \int_{S^{d-1}} \mathrm{d}\nu \int_{[\tau(\Phi),T]} \mathrm{d}\tilde{\tau}\, g\left(\Phi, \xi(\tilde{\Phi}, u + \tilde{\tau}(v-\tilde{v}) + a\nu)\right)\, \delta(t-\tilde{\tau})$$
$$= \int_{\mathcal{MT}} \mathrm{d}\lambda(\Phi) \int_{\mathcal{MT}_0} \mathrm{d}\lambda(\tilde{\Phi}) \int_{S^{d-1}} \mathrm{d}\nu\, g\left(\Phi, \xi(\tilde{\Phi}, u + t(v-\tilde{v}) + a\nu)\right).$$

Similarly we obtain for the right-hand side of (4.8)

$$\int_{\mathcal{MT}} \mathrm{d}\lambda(\Phi) \int_{S^{d-1}} \mathrm{d}\nu \int_{\mathcal{MT}} \mathrm{d}\lambda(\tilde{\Phi})\, g(\Phi,\tilde{\Phi})\, \delta\left(u(\Phi) - u(\tilde{\Phi}) + t(v(\Phi) - v(\tilde{\Phi})) - a\nu\right)$$
$$= \int_{\mathcal{MT}} \mathrm{d}\lambda(\Phi) \int_{S^{d-1}} \mathrm{d}\nu \int_{\mathcal{MT}_0} \mathrm{d}\lambda(\tilde{\Phi})\, g\left(\Phi, \xi\left(\tilde{\Phi}, u + t(v-\tilde{v}) - a\nu\right)\right).$$

Hence, both sides of the equation (4.8) coincide and the proof is finished. ☐

An immediate consequence of Lemma 4.1 is that the average of $\mathcal{Q}_t^a[P,P']$ is zero in the sense that

$$\int_{\mathcal{MT}} \mathrm{d}\mathcal{Q}_t^a[P,P'](\Phi)\, \varphi(\bar{\Phi}_t) = 0 \quad \text{for all } \varphi \in C_c(\mathcal{MT}). \tag{4.9}$$

The relation to the Boltzmann equation will become apparent in Proposition 4.8. To study the existence of solutions of (4.4), we first introduce the appropriate function spaces. We define spaces of integrable functions on $\mathcal{MT}$ with general weights

$$X_\ell := \{f \in M(\mathcal{MT}) \mid \|f\|_{X_\ell} < \infty \text{ and } f(\,\cdot \mid \tau = t) \in L^1(\mathcal{MT}_t) \quad \forall t \geq 0\} \tag{4.10}$$

with

$$\|f\|_{X_\ell} = \sup\left\{ \int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f|(\Phi)\, w(u+sv,s)\,(1+|v|)^\ell \;\middle|\; \int_{U\times[0,T]} \mathrm{d}u\,\mathrm{d}s\, w(u,s) = 1 \right\},$$

and let $X = X_1$. Note that X_ℓ is a Banach space but it is not a subset of $L^1(\mathcal{MT})$ because the τ-marginal can have concentrations.

Remark 4.2. *If $P \in X_\ell$ and the τ-marginal is in $L^1([0,T])$, then $P \in L^1(\mathcal{MT})$.*

To see that X_ℓ is a Banach space we suppose that f_m is a Cauchy sequence in X_ℓ, then $f_m \to f$ in $M(\mathcal{MT})$. Since the sequence f_m converges it is also tight. The absolute continuity follows from the disintegration theorem [DM78], which provides the existence of $\sigma \in M([0,T])$ and a family of measures $f_t = f(\,\cdot \mid \tau = t) \in L^1(\mathcal{MT})$ such that for all $g \in C(\mathcal{MT})$ the formula

$$\int_{\mathcal{MT}} g \,\mathrm{d}f = \int_0^T \mathrm{d}\sigma(t) \int_{\mathcal{MT}_t} \mathrm{d}f_t(\Phi)\, g(\Phi)$$

holds. Then for $E_t \subset \mathcal{MT}_t$ of measure zero $f_t(E_t) = 0$ for σ-almost every t, such that we have $f \in X_\ell$ after a modification on a set of measure zero.
After this preparation we can derive an existence and uniqueness result for the linearized evolution.

Lemma 4.3. *For each $P' \in C^0([0,T],X)$ the operator $\mathcal{Q}^a_t[\,\cdot\,, P'_t]$ is the generator of a strongly continuous evolution $U^a(s,t)$ on X, i.e. there exits a unique solution of the equation*

$$\frac{\partial}{\partial t} P^a_t = \mathcal{Q}^a_t[P^a_t, P'_t], \quad P_0 = f_0. \tag{4.11}$$

For each $t > 0$ the solution P^a_t has the following properties

1. *P^a_t has a density, i.e. $P_t \in L^1(\mathcal{MT})$.*
2. *P^a_t is nonnegative, i.e. $\int_\Omega \mathrm{d}P_t(\Phi) \geq 0$ for all $\Omega \subset \mathcal{MT}$ measurable.*
3. *P^a_t is normalized, i.e. $\int_{\mathcal{MT}} \mathrm{d}P^a_t(\Phi) = 1$.*
4. *The Lagrangian root marginal $\pi[P^a_t]$ which is defined by*
$$\int_{U\times\mathbb{R}^d} \mathrm{d}\pi[P](u,v)\, g(u,v) = \int_{\mathcal{MT}} \mathrm{d}P(\Phi)\, g(u(\Phi), v(\Phi))$$
is independent of a and t, i.e. $\pi[P^a_t] = f_0$.

Remark 4.4. *As a consequence of Lemma 4.3.4 we obtain the following formula for the collision rate which involves only f_0 but not P^a_t:*

$$\mu_t[P^a_t](u,v) = \mu_t(u,v) = \int_{S^{d-1}} \mathrm{d}\nu \int_{\mathbb{R}^d} \mathrm{d}v'\, f_0(u + t(v-v') + a\nu, v')\,(v-v')\cdot\nu. \tag{4.12}$$

Proof of Lemma 4.3. As the role of a is irrelevant for this proof it will not be shown. We show that $\mathcal{Q}_t[\,\cdot\,, P']$ generates an evolution on X with X_2 being a subset of the domain of the unbounded operator $\mathcal{Q}_t[\,\cdot\,, P'_t]$, for this we use general results of [Paz83, chap. 5]. The aim is to prove the existence of an evolution operator, which is the non-autonomous version of a semigroup. We study the following resolvent equation

$$A_\lambda f = g \tag{4.13}$$

with $A_\lambda f = \lambda f - \mathcal{Q}[f\;, P']$ for $\lambda > 0$ and $g \in X$. It suffices to establish the existence of nonnegative solutions f if $g \geq 0$. Indeed, for general g we can decompose f and g into positive and negative parts: $f = f_+ - f_-$, $g = g_+ - g_-$; if $Af_\pm = g_\pm$ then $A(f_+ - f_-) = g_+ - g_-$. We consider two separate cases depending on whether the time coincides with the final collision time of a tree.
Then for $t \neq \tau(\Phi)$, we obtain

$$\lambda f(\Phi) + \mu(\Phi) f(\Phi) = g(\Phi), \tag{4.14}$$

i.e.

$$f(\Phi) = \frac{1}{\lambda + \mu(\Phi)} g(\Phi). \tag{4.15}$$

For $t = \tau(\Phi)$ the solution to

$$(\lambda f - \mathcal{Q}[f,P'])(\Phi) = \lambda f(\Phi) + \mu(\Phi) f(\Phi) - f(\bar{\Phi}_t) P'(\Phi'_t)\,((v - v_{\mathrm{root}}(\Phi'))\cdot\nu)_+ = g(\Phi), \tag{4.16}$$

is given by

$$f(\Phi) = \frac{1}{\lambda + \mu(\Phi)} \left(g(\Phi) + g(\bar{\Phi}_t) \frac{P'(\Phi'_t)\,(v - v_{\mathrm{root}}(\Phi'))\cdot \nu_+}{\lambda + \mu_t(\bar{\Phi}_t)} \right). \tag{4.17}$$

The key observation in (4.15) and (4.17) is that f is non-negative for non-negative g and $\lambda > 0$. Hence then for nonnegative f

$$\|f\|_X = \sup \left\{ \int_{\mathcal{MT}} \mathrm{d}f(\Phi)\,(1+|v|)\,w(u+sv,s) \;\middle|\; \int_{U\times[0,T]} \mathrm{d}u\,\mathrm{d}s\,w(u,s) = 1 \right\}$$

(without the $|\,.\,|$–bars) implies with (4.13) and (4.9) that for each $w \in L^1(U\times[0,T])$ the equation

$$\begin{aligned}
&\int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}f(\Phi)\;(1+|v|)\,w(u+sv,s) \\
=\;& \frac{1}{\lambda}\int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}g(\Phi)\,(1+|v|)\,w(u+sv,s) \\
&+\frac{1}{\lambda}\int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}f(\bar{\Phi}_t)\,(1+|v|)\,w(u+sv,s) \int_{\mathcal{MT}} \mathrm{d}L_t[P'](\Phi') \\
=\;& \frac{1}{\lambda}\int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}g(\Phi)\,(1+|v|)\,w(u+sv,s)
\end{aligned}$$

holds. The first equation is due to (4.13), the second equation follows from (4.9). Nonpositive right-hand sides f can be treated analogously. Thus

$$\|f\|_X = \frac{1}{\lambda}\|g\|_X.$$

This shows that $Q_t[\,\cdot\,,P'_t]$ is a stable family of generators with exponential growth rate $\omega = 0$ and bound $M = 1$. Furthermore as (4.9) also holds when restricting to Φ with given root data, then also $f \in X_\ell$ if $g \in X_\ell$.
We will demonstrate now that for each $f \in X_2$ (cf. eqn. (4.10)), $t \in [0,T]$ and $P' \in X$

$$\|\mathcal{Q}_t[f,P']\| \le 2\kappa_d\,\|f\,\mu_t[P']\|. \tag{4.18}$$

Indeed, for fixed $t \in [0,T]$ one obtains that

$$\begin{aligned}
&\|\mathcal{Q}_t[f,P']\|_X \\
\le& \int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f(\Phi)|\,(1+|v|)\,w(u+sv,s) \\
&\times \int_{S^{d-1}} \mathrm{d}\nu \int_{\mathcal{MT}} \mathrm{d}P'(\Phi')\,\delta(u-u'+t\,(v-v')+a\nu)\,(v-v')\cdot\nu_+ \\
&+ \int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f|(\Phi)\,\mu_t[P'](\Phi)\,(1+|v|)\,w(u+sv,s)
\end{aligned}$$

It is immediate that the first term and the second term coincide:

$$\begin{aligned}
&\int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f(\Phi)|\,(1+|v|)\,w(u+sv,s) \\
&\times \int_{S^{d-1}} \mathrm{d}\nu \int_{\mathcal{MT}} \mathrm{d}P'(\Phi')\,\delta(u-u'+t\,(v-v')+a\nu)\,(v-v')\cdot\nu_+ \\
=& \int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f(\Phi)|\,(1+|v|)\,w(u+s\,v,s)\,\mu_t[P'](\Phi),
\end{aligned}$$

and thus inequality (4.18) is established. Hence, we obtain for each $t \in [0,T]$

$$
\begin{aligned}
&\int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f|(\bar\Phi)\,(1+|v|)\,w(u+sv,s) \int_{\mathcal{MT}} \mathrm{d}|L_t[P'](\Phi')| \\
\le& 2\int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f(\Phi)|\,(1+|v|)\,w(u+sv,s) \\
&\times \int_{\mathcal{MT}} \mathrm{d}P'_t(\Phi'')\,\delta(u-u''+t(v-v'')+a\nu)\,(v-v'')\cdot\nu_+ \\
\le& 2\int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f(\Phi)|\,(1+|v|)\,w(u+sv,s) \\
&\times \sup\left\{\int_{S^{d-1}} \mathrm{d}\nu \int_{\mathcal{MT}} \mathrm{d}P'_t(\Phi'')\,\tilde w(u-u''+t(v-v'')+a\nu)\,(v-v'')\cdot\nu_+ \;:\; \int_U \tilde w(u)\,\mathrm{d}u = 1\right\} \\
\le& 2\kappa_d \int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f|(\bar\Phi)\,(1+|v|)^2\,w(u+sv,s) \\
&\times \sup\left\{\int_0^T \mathrm{d}s \int_{\mathcal{MT}} P'_t(\Phi'')\,(1+|v''|)\,\tilde w(u''+sv'',s)) \;:\; \int_{U\times[0,T]} \mathrm{d}u\,\mathrm{d}s\,\tilde w(u,s) = 1\right\} \\
=& 2\kappa_d \int_0^T \mathrm{d}s \int_{\mathcal{MT}} \mathrm{d}|f|(\bar\Phi)\,(1+|v|)^2 w(u+sv,s)\,\|P'\|_X
\end{aligned}
$$

This implies that

$$\|\mathcal{Q}_t[f,P']\|_X \le 2\kappa_d\,\|f\|_{X_2}\|P'\|_X, \tag{4.19}$$

hence $X_2 \subset D(Q[\,\cdot\,,P'_t])$.
We are now in a position to check conditions $(H_1),(H_2),(H_3)$ in [Paz83, Theorem 5.3.1]. The first two conditions are satisfied as $Q[\,\cdot\,,P_t]$ is a stable family of generators with exponential growth rate $\omega = 0$ and bound $M = 1$ both on X and $Y = X_2$. By (4.19) and linearity in the second argument we also obtain that $t \mapsto Q[\,\cdot, P'_t]$ is continuous in the $\|\cdot\|_{Y\to X}$ norm as long as $t \mapsto P'_t$ is continuous in X. Then there exists a unique evolution system $U(t,t_0)$ by [Paz83, Theorem 5.3.1] on X.
Next we show that $P_t \ge 0$. The construction of $U(t,s)$ is based on repeated applications of $\mathrm{Id} - hQ[\cdot\,,P'_s]$ with $h > 0$ and $s \in [0,T]$, all these operators are multiples of the resolvent in (4.13). Hence they map positive functions to positive functions by the observation after (4.17). Thus we have $P_t \ge 0$.
Together with (4.9) this implies that for each $t \ge 0$ the measure P_t characterizes a probability distribution on $\mathcal{MT}$.
Next we show that the measure P_t has a density with respect to the Lebesgue measure if $t > 0$. Following Remark 4.2 it suffices to show that there exists a function $f \in L^1([0,\infty))$ with the property

$$\int_{\mathcal{MT}} \mathrm{d}P(\Phi)\,g(\tau(\Phi)) = \int_t \mathrm{d}f(t)\,g(t) \qquad \text{for all } g \in L^\infty([0,\infty)). \tag{4.20}$$

First note that (4.20) is an immediate consequence of the stronger bound

$$\int_{\mathcal{MT}} \mathrm{d}P(\Phi)\,g(\tau(\Phi)) \le C\|g\|_{L^1([0,T])} \qquad \text{for all } g \in L^1([0,\infty)).$$

As P_t solves (4.11) strongly, $\partial_t P \in C^0((0,T),X)$ and we have that for each $t \in [0,T]$

$$\int_{\mathcal{MT}} \mathrm{d}P_t(\Phi)\,g(\tau(\Phi)) \le \int_0^t \mathrm{d}\tau\,\mathrm{d}\mathcal{Q}^+_\tau(\Phi)\,g(\tau) \le \|g\|_{L^1([0,T])} \sup_{\tau\in[0,T]} \|\partial_t P(\tau)\|_X.$$

Finally part (iv) follows from (4.9) by letting $\phi(\bar\Phi_t) = g(u(\Phi_t),v(\Phi_t)) = g(u(\Phi),v(\Phi))$. □

We will also need an L^1 version of (4.19).

Remark 4.5. *Lemma 4.3 also holds if X is replaced with the Banach space*

$$Z = \{f \in L^1(\mathcal{MT}) \mid \|f\|_Z < \infty\}$$

with $\|f\|_Z = \int_{\mathcal{MT}} \mathrm{d}|f|(\Phi)\,(1+|v|)$ since estimate (4.19) follows with X replaced by Z, as $\mu_t(u,v)$ is an $L^1(\mathbb{T}^d)$ function for the argument $(u+tv)$, whereas $f \in X_2$ is $L^\infty(\mathbb{T}^d)$ for the same argument, such that the estimate of the product term $\mathcal{Q}_t[f, P']$ follows by the Hölder inequality. Note however that Z is not a suitable space for establishing the existence of a nonlinear semigroup.

Using a hierarchy of approximations we are now able to obtain the idealized distributions.

Proposition 4.6. (i) *For each $a > 0$ the nonlinear Kolmogorov equation (4.4) has a unique solution $P_t^a \in C^1([0,T], X) \cap C^0([0,T], X_2)$ for every f_0 satisfying (1.5).*
(ii) *For given initial data and for each t the measure P_t^a converges to $P_t = P_t^0$ in Z as $a \to 0$.*

Proof. As the role of a is not relevant for (i) we will not show the dependency on a in this part of the proof. We prove (i) by approximating P_t by a sequence of probability measures $P_{t,k}$ which are defined recursively by the equation

$$P_{t,1} = f_0, \tag{4.21}$$

$$\frac{\partial P_{t,k}}{\partial t} = \mathcal{Q}_t[P_{t,k}, P_{t,k-1}], \quad P_{0,k} = f_0, \tag{4.22}$$

The existence of evolution operator for (4.21, 4.22) if $P_{t,k} \in C^1([0,T], X) \cap C^0([0,T], Y)$ is a consequence of Lemma 4.3. To have classical solutions of the operator equation we have to use some more semigroup theory. The evolution system in Lemma 4.3 is constructed through an implicit Euler approximation, i.e. using a resolvent as in (4.13), then as the resolvents leave Y invariant, $U(t,t_0)$ maps Y and also any other X_ℓ to itself, giving condition (E_4) in [Paz83, Theorem 5.4.3].
To check the strong continuity condition in Y, condition (E_5), we start with initial data in $f \in X_2$, and use the previous results with X replaced by X_2 and Y by X_3. Then Theorem 5.3.1 in [Paz83] implies that there is a unique Y-valued solution of

$$\begin{cases} \frac{\partial P_t}{\partial t} &= \mathcal{Q}_t[P_t, P_t'], \\ P_0 &= f_0. \end{cases} \tag{4.23}$$

Replacing P' by $P_{t,k-1}$ and P_t by $P_{t,k}$ gives that $P_{t,k} \in C^1([0,T], X) \cap C^0([0,T], Y)$ for all $k \in \mathbb{N}$ by induction.
Next we will prove that $P_t = \lim_{k\to\infty} P_{t,k}$ exists by showing that the solutions of the non-autonomous linear equation (4.23) are a contraction of $P' \in C^0([0,T], X)$ with respect to the norm

$$\|P'\|_\rho = \sup_{t\in[0,T]} \exp(-\rho t)\|P_t'\|_X.$$

To consider a solution of (4.23), we replace P' with $\tilde{P}_t'$:

$$\begin{cases} \frac{\partial \tilde{P}_t}{\partial t} &= \mathcal{Q}_t[\tilde{P}_t, \tilde{P}_t'], \\ P_0 &= f_0, \end{cases} \tag{4.24}$$

then $P_t - \tilde{P}_t$ satisfies

$$\begin{cases} \frac{\partial P_t}{\partial t} - \frac{\partial \tilde{P}_t}{\partial t} &= \mathcal{Q}_t[P_t - \tilde{P}_t, \tilde{P}_t'] + \mathcal{Q}_t[P_t, P_t' - \tilde{P}_t'], \\ P_0 - P_0 &= 0. \end{cases} \tag{4.25}$$

Then the strong solution $P_t - \tilde{P}_t \in C^1([0,T], X) \cap C^0([0,T], X_2)$ constructed above can represented as a mild solution of (4.25). Lemma 4.3 gives the evolution $U(t,s)$ generated by P', thus

$$P_t - \tilde{P}_t = \int_0^t U(t,s)\,\mathcal{Q}_s[P_s, P_s' - \tilde{P}_s']\,\mathrm{d}s.$$

Then (4.19), the boundedness of $U(t,s)$ and the fact that $\|P_s\|_{X_2} \le K_\infty$ for all $s \ge 0$ gives the estimate

$$\begin{aligned}\|P - \tilde{P}\|_\rho &= \sup_{t\in[0,T]} \exp(-\rho t) \left\| \int_0^t U(t,s)\, \mathcal{Q}_s[P_s, P_s' - \tilde{P}_s'] \,\mathrm{d}s \right\|_X \\ &\le 2\kappa_d \sup_{t\in[0,T]} \exp(-\rho t) \int_0^t \|P_s\|_{X_2} \, \|P_s' - \tilde{P}_s'\|_X \,\mathrm{d}s \\ &\le 2\kappa_d \sup_{t\in[0,T]} \exp(-\rho t) \int_0^t K_\infty \exp(\rho s)\, \|P' - \tilde{P}'\|_\rho \,\mathrm{d}s \\ &\le \frac{2\kappa_d K_\infty}{\rho} \left(\sup_{t\in[0,T]} (1 - \exp(-\rho t)) \right) \|P' - \tilde{P}'\|_\rho,\end{aligned}$$

i.e. this is a contraction for $\rho > 2\kappa_d K_\infty$. Thus $P_{\cdot,k}$ converges in $C^0([0,T], X)$ to a unique fixed point P. Setting $P_t' = P_t$ in (4.23) and using Lemma 4.3 then gives the desired regularity.
To prove (ii), we reintroduce the parameter a with the convention that $P_t = P_t^0$. The difference $P_t^a - P_t$ satisfies

$$\frac{\partial P_t^a}{\partial t} - \frac{\partial P_t}{\partial t} = \mathcal{Q}_t[P_t^a - P_t, P_t] + \mathcal{Q}_t^a[P_t^a, P_t^a - P_t] + P_t^a(L_t^a[P_t] - L_t[P_t]).$$

Denoting the evolution generated by $\mathcal{Q}_t[\,\cdot\,, P_t]$ as $U(t,s)$, we obtain

$$P_t^a - P_t = \int_0^t U(t,s) \left\{ \mathcal{Q}_s^a[P_s^a, P_s^a - P_s] + P_s^a(L_s^a[P_s] - L_s[P_s]) \right\} \mathrm{d}s.$$

By Remark 4.5 we obtain the bound

$$\|\mathcal{Q}^a[P, P']\|_Z \le 2\kappa_d \, \|P\|_{X_2} \|P_t'\|_Z. \tag{4.26}$$

Using (4.26) and that $U(t,s)$ is a bounded operator on Z we arrive at

$$\begin{aligned}&\|P_t^a - P_t\|_Z \\ =& \int_0^t \|U(t,s)\, \mathcal{Q}_s^a[P_s^a, P_s^a - P_s]\|_Z \,\mathrm{d}s + \int_0^t \|P_s^a(L_s^a[P_s] - L_s[P_s])\|_Z \,\mathrm{d}s \\ \le& C \int_0^t \left\{ \|P_s^a\|_{X_2} \|P_s^a - P_s\|_Z + \|P_s^a(L_s^a[P_s] - L_s[P_s])\|_Z \right\} \mathrm{d}s\end{aligned}$$

Due to strong continuity of spatial shift in L^1 norm used for Z the last term converges to 0 as $a \to 0$. Gronwall's inequality gives the required convergence in (ii). ☐

4.2. **Properties of the idealized distribution.** The existence result Proposition 4.6(i) delivers a tightness bound on the number of nodes in P_t of the form: There exists a function $M(\varepsilon)$ such that

$$P_t(\{\#\Phi \ge M(\varepsilon)\}) \le \varepsilon \tag{4.27}$$

for all $\varepsilon > 0$.
We will show next that it is possible to find marginal distributions which satisfy closed evolution equations. The fact that P_t itself is a Markovian process is trivial since P_t is historical, i.e. for each tree $\Phi \in \mathcal{MT}$ and $t \ge 0$ the past trees Φ_s are still visible for Φ_t, cf. [DP91]. The complexity of the tree increases monotonously in time, i.e. information does not get lost. On the other hand, it is possible to find Markovian random variables with constant complexity. As an intermediate step towards constructing Markovian random variables we show that the subtrees with collision times $t \ge s$ are Markovian with respect to $(u(s), v(s))$.

Definition 4.7. *For each tree $\Phi \in \mathcal{MT}$ the random variable $\beta(\Phi) \in \{0,1\}$ is defined recursively by*

$$\beta(\Phi) = \begin{cases} 1 & \text{if } \#\Phi = 1, \\ \beta(\bar{\Phi}_\tau)(1 - \beta(\Phi_\tau')) & \text{else.} \end{cases} \tag{4.28}$$

The random variable $\beta(\Phi)$ is the indicator function of those trees where the root particle has not undergone a collision. We will show now that the expectation of this observable satisfies a closed evolution equations.

Proposition 4.8. *Consider $P_t = P_t^0$ as in Proposition 4.6. The marginal distribution*

$$f_t(u,v) = P_t(\beta = 1 \text{ and } (u_{\mathrm{root}}, v_{\mathrm{root}}) = (u - tv, v))$$

satisfies the closed equations and initial conditions

$$\frac{\partial f}{\partial t} + v \cdot \nabla_u f = -Q_-[f,f], \qquad f_{t=0}(u,v) = f_0(u,v), \tag{4.29}$$

in the weak sense, where $Q_-[\,\cdot\,,\,\cdot\,]$ is defined in (2.6).

The appearance of the transport term $v \cdot \nabla_u f$ is a result of the change from Lagrangian to Eulerian coordinates: P_t provides the distribution of the initial positions and velocities and $f_t(u,v)$ characterize the densities of particles with velocity v at position u at time t. An analogous statement is also true if $a > 0$, but this is irrelevant for our purposes.

Proof. Let $g \in C^1(U \times \mathbb{R}^d)$ be a testfunction. Then

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}f_t(u,v)\, g(u,v) = \frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}P_t(\Phi)\, \beta(\Phi)\, g(u+tv, v) \\
&= \int_{\mathcal{MT}} \mathrm{d}\mathcal{Q}_t[P_t, P_t](\Phi)\, \beta(\Phi)\, g(u+tv,v) + \int \mathrm{d}P_t(\Phi)\, \beta(\Phi)\, v \cdot \nabla g(u+tv, v) \\
&= \int_{\mathcal{MT}} \mathrm{d}\mathcal{Q}_t[P_t, P_t](\Phi)\, \beta(\Phi)\, g(u+tv,v) + \int_{U \times \mathbb{R}^d} \mathrm{d}f_t(u+tv, v)\, v \cdot \nabla g(u+tv, v) \\
&= I_1 + I_2.
\end{aligned}$$

A change of variables in I_2 yields that

$$I_2 = \int_{U \times \mathbb{R}^d} \mathrm{d}f_t(u,v)\, v \cdot \nabla g(u,v).$$

We analyze now I_1. The definition of β implies that

$$\begin{aligned}
I_1 &= \int_{\mathcal{MT}} \mathrm{d}\mathcal{Q}_t[P_t, P_t](\Phi)\, \beta(\bar{\Phi}_t)\, g(u+tv, v)\, \mathbf{1}_{\tau < t} \\
&\quad + \int_{\mathcal{MT}_t} \mathrm{d}\mathcal{Q}_t[P_t, P_t](\Phi)\, \beta(\Phi)\, g(u+tv,v)\, \beta(\bar{\Phi}_t)\, (1 - \beta(\Phi_t')) \\
&= \int_{\mathcal{MT}} \mathrm{d}\mathcal{Q}_t[P_t, P_t](\Phi)\, \beta(\bar{\Phi}_t)\, g(u+tv, v) \\
&\quad - \int_{\mathcal{MT}_t} \mathrm{d}\mathcal{Q}_t[P_t, P_t](\Phi)\, g(u+tv,v)\, \beta(\bar{\Phi}_t)\, \beta(\Phi_t').
\end{aligned}$$

Thanks to formula (4.9) the first term vanishes. The definition of $\mathcal{Q}_t$ implies that

$$\begin{aligned}
&\int_{\mathcal{MT}_t} \mathrm{d}\mathcal{Q}_t[P_t, P_t](\Phi)\, \beta(\bar{\Phi}_t)\, \beta(\Phi_t')\, g(u+tv, v) \\
&= \int_{U \times \mathbb{R}^d} \mathrm{d}Q_-[f_t, f_t](u+tv, v)\, g(u+tv, v) = \int_{U \times \mathbb{R}^d} \mathrm{d}Q_-[f_t, f_t](u,v)\, g(u,v).
\end{aligned}$$

The last equality follows from a change of variables. Putting everything back together we find that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}f_t(u,v)\, g(u,v) \\
&= -\int_{U \times \mathbb{R}^d} \mathrm{d}g(u,v)\, Q_-[f_t, f_t](u,v) + \int_{U \times \mathbb{R}^d} \mathrm{d}f_t(u,v)\, v \cdot \nabla_u g(u,v).
\end{aligned}$$

for all testfunctions $g \in C^1(U \times \mathbb{R}^d)$, which is the claim. □

The link between the mild solutions in Proposition 4.6 and weak solutions in Proposition 4.8 is provided by the following proposition. We derive a formula that can be evaluated for a wide class of measures.

Proposition 4.9. *Let* $f_t \in L^2(U, L^1(\mathbb{R}^d))$.

(i) *The equation*

$$f_t(u,v) = f_0(u - tv, v) \exp\left(- \int_0^t \int_{\mathbb{R}^d} f_s(u - (t-s)v, v')\, \kappa_d\, |v - v'| \, \mathrm{d}v' \, \mathrm{d}s \right), \tag{4.30}$$

is satisfied for all $t \in (0,T)$, *if and only if* f_t *is the unique mild solution of (2.5).*

(ii) *Equation (4.30) implies that* f_t *is also a distributional solution. Furthermore, every distributional solution with* $Q_-[f,f] \in L^1\left((0,T) \times U, L^1_{1+|v|}(\mathbb{R}^d)\right)$ *is a mild solution.*

Proof. An equivalent formulation for f_t being a mild solution (2.7) is

$$f^\#(u,v,t) - f_0(u,v) = - \int_0^t (Q_-[f_s, f_s])^\# \mathrm{d}s \tag{4.31}$$

with $h^\#(u,v,t) = h(u + tv, v, t)$.
First we show (i). Let f_t be a mild solution, then

$$\begin{aligned} f^\#(u,v,t) - f_0(u,v) &= - \int_0^t (L[f_s] f_s)^\# \mathrm{d}s = - \int_0^t (L[f_s])^\# f_s^\# \mathrm{d}s \\ &= - \int_0^t g(u,v,s) f_s^\# \mathrm{d}s. \end{aligned} \tag{4.32}$$

with $g(u,v,s) = (L[f_s])^\#(u,v) = \int_{\mathbb{R}^d} f_s(u + sv, v') \kappa_d |v - v'| \, \mathrm{d}v'$ such that $g(u, \cdot, s) \frac{1}{1+|\cdot|} \in L^2_{loc}(\mathbb{R}^d)$. A solution to (4.32) is given by

$$f^\#(u,v,t) = \exp\left(- \int_0^t g(u,v,s) \, \mathrm{d}s \right) f_0(u,v), \tag{4.33}$$

as for each u, v the equation decouples to a single scalar ordinary differential equation, so f_t fulfills (4.30). For fixed u, the integral equation (4.33) has a unique solution in $C^0([0,T], L^1_{1+|v|}(\mathbb{R}^d))$ by a simple contraction argument for finite times as in the spatially homogeneous case [MT10, Lemma 5]. This observation also shows that f_t given by (4.30) is a mild solution, completing the proof of (i).
For part (ii), we observe that mild and distributional solutions coincide following [Ba77], for

$$\partial_t f + v \cdot \nabla_u f = h$$

as long as $h = -Q_-[f_t, f_t] \in L^1((0,T), L^1_{1+|v|}(U \times \mathbb{R}^d))$, which immediately shows the second part of (ii). For solutions given by (4.30), we first observe $0 \le f_t(u,v) \le f_0(u - tv, v)$ by (4.33). Hence

$$0 \le Q_-[f_t, f_t] = L[f_t](u,v) f_t(u,v) \le L[f_0(u - tv, \cdot)](v) f_0(u - tv, v),$$

such that $Q_-[f_t, f_t] \in L^1_{1+|v|}(U \times \mathbb{R}^d)$ as

$$\|L[f_0 \circ \varphi] \, f_0 \circ \varphi\|_{L^1_{1+|v|}} = \|L[f_0] \, f_0\|_{L^1_{1+|v|}} < \infty$$

with $\varphi(u,v) = (u - tv, v)$. The last equation is due to (1.5), this completes the proof. □

Remark 4.10. *Interestingly there exists an explicit solution of the nonlinear Kolmogorov equation (4.4), but this fact is not relevant for our analysis. Let* $\Omega \subset \mathcal{MT}$ *and* $t \in [0, \infty)$. *Then the idealized distribution is given by*

$$P_t(\Omega) = \int_\Omega \exp\Big(-I_t(\Phi) \Big) \, \mathrm{d}\lambda(\Phi) \tag{4.34}$$

where the integrated collision rate is I_t is defined recursively

$$I_t(\Phi) = \int_0^t \Gamma_s(\Phi)\,\mathrm{d}s + \sum_{s\in[0,t]} I_s(\Phi'_s), \tag{4.35}$$

$$\Gamma_t(\Phi) = \int_{\mathbb{R}^d} \kappa_d f_0(u - t(v-v'), v')\,(v-v')_+\,\mathrm{d}v', \tag{4.36}$$

$$\mathrm{d}\lambda_t(\Phi) = f_0(u,v)\,\mathrm{d}u\,\mathrm{d}v \times \begin{cases} 1 & \text{if } \Phi = \emptyset, \\ \prod_{s\in[0,t]} \kappa_d \lambda_s(\Phi'_s)(v - v'_s)_+ & \text{else.} \end{cases} \tag{4.37}$$

Note that the convention $I_t(\emptyset) = 0$ is used. The initial positions $u_l \in U$ are defined by the formula (3.1) with $a = 0$. For $a > 0$ an analogous formula can be obtained.

We end the section by introducing trees without recollisions.

Definition 4.11. *A tree Φ is recollision free on the time interval $[0,t]$ at diameter a if*

$$|u_l^a + sv_l - u_{l'}^a - sv_l| > a \text{ for all } 0 < s < t \text{ and } l, l' \in m(\Phi) \text{ such that } l \not\sim l'. \tag{4.38}$$

For any pair of monotonic functions $M(a)$, $V(a)$ such that $\lim_{a\to 0} M(a) = \lim_{a\to 0} V(a) = +\infty$ the set of good trees is defined as

$$\mathcal{G}(a_0) = \Big\{ \Phi \in \mathcal{MT} \;:\; \#\Phi \le M(a_0) \text{ and } \max_{l\in m(\Phi)} |v_l| \le V(a_0) \text{ and } \min_{l \in m(\Phi)\setminus \mathrm{root}} \nu_l \cdot (v_l - v_{\bar{l}}) > 0 \text{ and (4.38) holds for all } a \in [0, a_0] \Big\}. \tag{4.39}$$

Note that thanks to the monotonicity of V and M the set $\mathcal{G}$ is monotonic in a_0.

Lemma 4.12. *The set of good trees has almost full measure, i.e.*

$$\lim_{a\to 0} P_t^a(\mathcal{G}(a)) = \lim_{a\to 0} P_t(\mathcal{G}(a)) = 1. \tag{4.40}$$

Proof. We first show that $\mathcal{G}(0)$ is a set of measure 1. The only nontrivial condition is (4.38) with $a = 0$. Let $\sigma \in \mathcal{T}$ be a tree skeleton and define $\mathcal{MT}(\sigma) := \{\Phi \in \mathcal{MT} \;:\; m(\Phi) = \sigma\}$. Recall equation (3.1) which provides for each $l \in \sigma$ a recursive formula for the initial position u_l. We will write $u_{l'}(s_{l'}, v_{l'})$ to emphasize the dependency of the the initial position on the collision time $s_{l'}$ and velocity $v_{l'}$.
The dimension of $\mathcal{MT}(\sigma)$ is $(2d)^{\#\sigma}$ as the nodes are parameterized by $(u,v) \in \mathbb{T}^d \times \mathbb{R}^d$ for the root and by $(v, \nu, s) \in \mathbb{R}^d \times S^{d-1} \times [0,T]$. On the other hand, for $a = 0$ any pair $l, l' \in \sigma$ and fixed $(u_l, v_l) \in U \times \mathbb{R}^d$ and fixed $(u_{\bar{l}'}, v_{\bar{l}'}) \in U \times \mathbb{R}^d$, the subset of $\mathcal{MT}(\sigma)$ with

$$\{(v_{l'}, \nu_{l'}, s_{l'} \in \mathbb{R}^d \times S^{d-1} \times [0,T] \;:\; u_l - u_{l'}(s_{l'}, v_{l'}) = t(v_{l'} - v_l) \text{ for some } t \in [0,T]\} \tag{4.41}$$

is of zero measure by a simple dimension argument. To see this, we first express $u_{l'} = u_{\bar{l}'} + s_{l'}(v_{\bar{l}'} - v_{l'})$ by (3.1). Then the condition in (4.41) is

$$(t - s_{l'})v_{l'} = u_l - u_{\bar{l}'} - s_{l'}v_{\bar{l}'} + tv_l,$$

for given $t \neq s_{l'}$ the velocity $v_{l'}$ is the countable set $v_{l'} = \frac{1}{t - s_{l'}}\left(u_l - u_{\bar{l}'} - s_{l'}v_{\bar{l}'} + tv_l + \mathbb{Z}^d\right)$ giving restriction to a collection of $2 + (d-1)$ dimensional sets. If on the other hand $t = s_{l'}$, then (4.41) gives $u_l - u_{\bar{l}'} + t(-v_{\bar{l}'} + v_l) = 0$. This implies that $v_{l'}$ is arbitrary but there at most a finite number of $t \in [0,T]$ satisfying this equality in $\mathbb{T}^d$, these subsets of $v_{l'}, \nu_{l'}, s_{l'}$ are hence $2d - 1$ dimensional.
Thus $\mathcal{MT}(\sigma)\setminus\mathcal{G}(0)$ is a countable union of manifolds of dimension less or equal to $(2d)^{\#\sigma} - 1$ Since $\mathcal{MT} = \cup_{\sigma\in\mathcal{T}} \mathcal{MT}(\sigma)$ and $P_t \in L^1(\mathcal{MT})$ we obtain that $P_t(\mathcal{MT} \setminus \mathcal{G}(0)) = 0$ and $P_t(\mathcal{G}(0)) = 1$.
Furthermore, for each $\Phi \in \mathcal{G}(0)$ there exists a_0 such that $\Phi \in \mathcal{G}(a)$ for all $a < a_0$. Hence $\lim_{a\to 0} \mathcal{G}(a) = \mathcal{G}(0)$ and dominated convergence implies $\lim_{a\to 0} P_t(\mathcal{G}(a)) = P_t(\mathcal{G}(0)) = 1$. Thanks to the convergence in Proposition 4.6(ii) we obtain the remaining claim $\lim_{a\to 0} P_t^a(\mathcal{G}(a)) = 1$. □

Remark 4.13. *For $\alpha > 1$ and $\Phi \in \mathcal{MT}^\alpha$ we define*

$$P_t^\alpha(\Phi) = \prod_{i=1}^{\alpha} P_t(\Phi^{(i)}),$$

and accordingly for $a > 0$.

5. Empiric distribution

Now we consider the empiric distribution $\hat{P}_t$ defined by the Newton dynamics (2.3) and the scattering state β^a (2.4). The initial values of the particles form a random set $\omega \subset U \times \mathbb{R}^d$. As already indicated at the beginning of the section, the family of probability measures $\hat{P}_t \in PM(\mathcal{MT}^1)$ is the marginal distribution of the tree Φ which is generated by the many-body evolution and has the first particle as its root. Similarly, we can consider trees with several root particles and analyze trees in $\mathcal{MT}^\alpha$. The key result in this section is the demonstration that the empiric distribution $\hat{P}_t$ satisfies a differential equation (5.4) which is very similar to the idealized equation (4.4). The main difference is given by factors like γ which capture dilution effects. The similarity of equations (5.4) and (4.4) enables us show later that the total variation distance between $\hat{P}_t$ and P_t converges to 0 as a tends to 0.
For every $\Omega \subset \mathcal{MT}^1$ measurable we define

$$\hat{P}_t(\Omega) := \mathrm{Prob}_{a,t}(\Phi^{(1)} \in \Omega). \tag{5.1}$$

Similar, for $\Psi \in \mathcal{MT}^\alpha$ the conditional distribution of Φ given $\Psi \in \mathcal{MT}^\alpha$ is defined as

$$\hat{P}_t(\Omega \mid \Psi) = \mathrm{Prob}_{a,t}\left(\Phi^{(1)} \in \Omega \;\middle|\; (\Phi^{(2)}, \ldots, \Phi^{(\alpha+1)}) = \Psi\right),$$

with the convention that $\hat{P}_t(\cdot \mid \Psi) = \hat{P}_t(\cdot)$ if $\Psi \in \mathcal{MT}^0$. We define the operator

$$\hat{\mathcal{Q}}_t^a : \mathrm{PM}(\mathcal{MT}) \times \mathrm{PM}(\mathcal{MT}) \times \mathcal{MT} \to M(\mathcal{MT})$$

by

$$\hat{\mathcal{Q}}_t^a[P, P'](\Phi) = P(\bar{\Phi}_t)\, \hat{L}_t^a[P'(\,\cdot \mid \bar{\Phi}_t)](\Phi)$$

where

$$\hat{L}_t^a[P](\Phi) = \delta(t - \tau(\Phi))\, P(\Phi_t')\, [(v(\Phi_t') - v(\Phi)) \cdot \nu_t]_+ - \hat{\mu}_t[P](\Phi). \tag{5.2}$$

The empiric collision rate $\hat{\mu}_t$ is defined by the expression

$$\hat{\mu}_t[P](\Phi) = \frac{1}{\int_{S^{d-1}} \mathrm{d}\nu \int_{\mathbb{R}^d} \mathrm{d}v'\, f_0(z_t)\, \mathbf{1}_{\Phi,t}(z_t)} \int_{S^{d-1}} \mathrm{d}\nu \int_{\mathbb{R}^d} \mathrm{d}v'\, f_0(z_t)\, \mathbf{1}_{\Phi,t}(z_t)\, [(v - v') \cdot \nu]_+,$$

$$\mathbf{1}_{\Psi,t}(u', v') = \begin{cases} 1 & \text{if } \min\{|u' - u_l + s(v' - v_l)| \;:\; s \in [0,t], l \in m(\Psi) \setminus \mathrm{root}\} > a, \\ 0 & \text{else.} \end{cases}$$

with the convention

$$z_t = (u + t(v - v') + a\nu, v). \tag{5.3}$$

Proposition 5.1. *Let $\alpha \geq 0$ and $\Gamma \in \mathcal{MT}^\alpha$. The empiric distribution $\hat{P}_t$ satisfies for all $\Phi \in \mathcal{G}$ the following differential equation*

$$\partial_t \hat{P}_t(\Phi \mid \Gamma) = (1 - \gamma)\, \hat{\mathcal{Q}}_t[\hat{P}_t(\,\cdot \mid \Gamma), \hat{P}_t(\,\cdot \mid \Gamma)](\Phi), \tag{5.4}$$

where $\gamma = (\#\bar{\Phi}_t + \#\Gamma)a^{d-1}$.

The proof of Proposition 5.1 relies on the a-priori information that $\hat{P}_t$ is absolutely continuous with respect to the Lebesgue measure.

Lemma 5.2. *Let $a > 0$ and $\Psi \in \mathcal{G}(a)$. The empiric distribution $\hat{P}_t$ is absolutely continuous with respect to the Lebesgue-measure λ in a neighborhood of Ψ. Moreover, for all $A \subset U \times \mathbb{R}^d$ measurable*

$$\hat{P}_t\left(z^{(1)}_{\text{root}}, z^{(2)}_{\text{root}} \in A \;\middle|\; \Psi \right) = \left(\frac{\int_{S^{d-1}} \mathrm{d}\nu \int_{\mathbb{R}^d} \mathrm{d}v' \, f_0(z_t)\, \mathbf{1}_A(z_t)\, \mathbf{1}_{\Psi,t}(z_t)}{\int_{S^{d-1}} \mathrm{d}\nu \int_{\mathbb{R}^d} \mathrm{d}v' \, f_0(z_t)\, \mathbf{1}_{\Psi,t}(z_t)} \right)^2, \tag{5.5}$$

with z_t given by (5.3).

Proof. Let $\omega \subset U \times \mathbb{R}^d$ be the set of initial positions and velocities. Observe first that the inequality

$$\hat{P}_t(\Psi) \leq \text{Prob}_a \left(\cup_{l \in m(\Psi)} \{\varphi_l(\Psi)\} \subset \omega \right),$$

holds because a good tree can only exist if the required initial states form a subset of ω. Next, note that the definition of $\mathcal{G}(a)$ implies that there exists $h > 0$ such that $B_h(\Psi) \subset \mathcal{G}(a)$. Let $\varphi : B_h(\Psi) \to (U \times \mathbb{R}^d)^{\# m(\Psi)}$ denote the initial positions and velocities of the particles in the tree. We will show next that

$$\det(\nabla \varphi(\Psi)) = \prod_{l \in m(\Psi) \setminus \text{root}} [a^{d-1}(v_l - v_{\bar{l}}) \cdot \nu_l]. \tag{5.6}$$

Formula (5.6) is a direct consequence of the observation that the gradient of φ with respect to (t_l, ν_l) is given by a simple expression. Indeed, let $F(\Psi) = \nabla \varphi(\Psi)$ be the gradient and define the submatrices $F_{lk} = \nabla_{t_l, \nu_l} \varphi_k(\Phi) \in \mathbb{R}^{d \times d}$ if $l \neq$ root and $F_{lk} = \nabla_{u_l} \varphi_k(\Phi)$ if $l =$ root. If the tree Ψ is generic (i.e. $s_l \neq s_k$ for all $l, k \in m(\Psi)$ such that $\bar{l} = \bar{k}$, cf Remark 3.2) then one obtains, in a coordinate system where $\nu_l = e_1 = (1, 0, \ldots, 0)^T$, that

$$F_{lk} = \begin{cases} \begin{pmatrix} (v_l - v_{\bar{l}}) \cdot \nu_l & 0 \\ * & a \, \text{Id}(d-1) \end{pmatrix} & \text{if } k = l \neq \text{root}, \\ \text{Id}(d) & \text{if } k = l = \text{root}, \\ 0 & \text{if } k > l, \end{cases} \tag{5.7}$$

where $\text{Id}(d)$ is the d-dimensional identity matrix and '$>$' denotes the semi-order on the set of set of vertices, cf. Definition 3.1. The ordering of the vertices implies that F is a block-triangular matrix. Hence the determinant of F is just the product of the determinants of F_{ll}, this yields equation (5.6).
Now we consider the cubes $C_{h,l} \subset U \times \mathbb{R}^d$ centered at $\varphi_l(\Psi)$ with sidelength h. As Ψ is good, the cubes are disjoint for sufficiently small h. If a and h are small the assumption that the initial values $z_1 \ldots z_n \in U \times \mathbb{R}^d$ are iid random variables with law f_0 and the scaling law (1.1) implies that

$$\text{Prob}_a(\#(\omega \cap C_{h,l}) = 1) = a^{1-d} f_0(C_{h,l})(1 + o(1)) \text{ as } h \to 0 \tag{5.8}$$

for each $l \in m(\Psi)$ where by a slight abuse of notation we use f_0 as a measure, i.e. $f_0(C_{h,l}) = \int_{C_{h,l}} f_0(u, v) \, \mathrm{d}u \, \mathrm{d}v$. Let now $C_h(\Psi) = \prod_{l \in m(\Psi)} C_{h,l}(\Psi)$. The formula

$$\lambda \left(\varphi^{-1} \left(C_h \right) \right) = h^{2d \# \Psi} \det(F(\Psi))^{-1} (1 + o(1)) \text{ as } h \to 0$$

together with equation (5.6) implies that

$$\frac{\hat{P}_t(\varphi^{-1}(C_h))}{\lambda(\varphi^{-1}(C_h))} = \frac{f_0(C_{h,\text{root}})}{h^{2d}} \prod_{l \in m(\Psi) \setminus \text{root}} \left[\frac{f_0(C_{h,l})}{h^{2d}(1 + o(1))} (v_l - v_{\bar{l}}) \cdot \nu_l \right]. \tag{5.9}$$

Since $f_0 \in L^1(U \times \mathbb{R}^d)$, the right-hand side in (5.9) remains bounded as $h \to 0$ for almost every $\Psi \in \mathcal{G}(a)$. This establishes the absolute continuity of $\hat{P}_t$. Formula (5.5) is an immediate consequence of the definition of $\hat{P}(\cdot \mid \Psi)$ and the assumption that $z^{(1)}, z^{(2)}, \ldots$ are iid random variables with law f_0. The proof of the lemma is complete. $\square$

Proof of Proposition 5.1. To simplify the notation we assume that $\Gamma = \emptyset$, which means that no conditioning is active. The general case is analogous as explained at the end of the proof. First consider $\Phi \in \mathcal{MT}$ with the property $\tau(\Phi) = t$. Recall that ν_t denotes the impact parameter of the final collision.

The probability $\hat{P}_t(\Psi)$ can be expressed in terms of the probabilities of Ψ_t' and $\bar{\Psi}_t$:

$$\begin{aligned}\hat{P}_t(\Psi) =&\hat{P}_t\left(\Phi^{(1)} = \Psi\right) = \hat{P}_t\left(\Phi_t^{(1)'} = \Psi_t' \text{ and } \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right)\\ =&\hat{P}_t\left(\Phi_t^{(1)'} = \Psi_t' \;\middle|\; \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right)\, \hat{P}_t\left(\bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right).\end{aligned}$$

The key idea is that the first probability can be expressed in terms of a two-body event. We will now demonstrate that

$$\hat{P}_t\left(\Phi_t^{(1)'} = \Psi_t' \;\middle|\; \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right) = (1-\gamma)\, \hat{P}_t\left(\Phi_t^{(1)} = \Psi_t' \;\middle|\; \Phi_t^{(2)} = \bar{\Psi}_t\right)\, [(v-v')\cdot\nu_t]_+ \tag{5.10}$$

by establishing matching upper and lower bounds. For each $h > 0$ define the set

$$\begin{aligned} U_h &= \{\Phi \in B_h(\Psi) \;:\; \Phi_l = \Psi_l \text{ for all } l \in m(\bar{\Psi}_t)\},\\ V_h &= \{\Phi_t' \;:\; \Phi \in U_h\}.\end{aligned}$$

The absolute continuity of $\hat{P}_t$ (Lemma 5.2) implies that

$$\hat{P}_t\left(\Psi_t' \;\middle|\; \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right) = \lim_{h\to 0} h^{-2d}\hat{P}_t\left(\Phi^{(1)} \in U_h \;\middle|\; \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right).$$

Since there are at most n possible choices for the index of the colliding particle we find that

$$\hat{P}_t\left(\Phi^{(1)} \in U_h \;\middle|\; \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right) \le \sum_{i=1}^{n} \hat{P}_t\left(\Phi^{(i)} \in V_h \;\middle|\; \Phi^{(1)} = \bar{\Psi}_t\right). \tag{5.11}$$

The permutation invariance and the fact that the particles in $\bar{\Psi}_t$ are ruled out as collision partners implies that

$$\sum_{i=1}^{n} \hat{P}_t\left(\Phi^{(i)} \in V_h \mid \Phi^{(1)} = \bar{\Psi}_t\right) =(n - \#\bar{\Psi}_t)\hat{P}_t\left(\Phi^{(1)} \in V_h \mid \Phi^{(2)} = \bar{\Psi}_t\right).$$

Formula (5.7) implies that

$$\lim_{h\to 0} h^{-2d}\hat{P}_t\left(\Phi^{(1)} \in V_h \mid \Phi^{(2)} = \bar{\Psi}_t\right) =P_t\left(\Phi_1 = \Psi_t' \mid \Phi_2 = \bar{\Psi}_t\right) a^{d-1}[(v-v')\cdot\nu_t]_+,$$

and thereby delivers the upper bound

$$\hat{P}_t\left(\Phi^{(1)} = \Psi_t' \mid \bar{\Phi}_t^{(2)} = \bar{\Psi}_t\right) \le (1-\gamma)\, \hat{P}_t\left(\Phi^{(1)} = \Psi_t' \mid \Phi^{(2)} = \bar{\Psi}_t\right) [(v-v')\cdot\nu_t]_+. \tag{5.12}$$

Next we derive the corresponding lower bound

$$\hat{P}_t\left(\Phi_t^{(1)'} = \Psi_t' \mid \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right) \ge (1-\gamma)\, \hat{P}_t\left(\Phi^{(1)} = \Psi_t' \mid \Phi^{(2)} = \bar{\Psi}_t\right) [(v-v')\cdot\nu_t]_+. \tag{5.13}$$

Define the set of initial values leading to a collision within the time interval $[t-h, t]$:

$$W_h = \left\{(u', v') \in U \times \mathbb{R}^d \;:\; \min_{t-h\le s\le t} |u - u' + s(v-v')| \le a\right\}.$$

Clearly

$$\begin{aligned}\hat{P}_t\left(\Phi^{(1)} \in U_h \;\middle|\; \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right) \ge &\sum_{i=1}^{n} \hat{P}_t\left(\Phi^{(i)} \in V_h \;\middle|\; \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right)\\ &- \operatorname{Prob}\left(\#(\omega \cap W_h) \ge 2 \mid \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right),\end{aligned}$$

and using the inclusion-exclusion principle we obtain that

$$\operatorname{Prob}\left(\#(\omega \subset W_h) \ge 2 \mid \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right) \le \sum_{1\le i<j\le n} \hat{P}_t\left(\{z^{(i)}, z^{(j)}\} \subset W_h \;\middle|\; \bar{\Phi}_t^{(1)} = \bar{\Psi}_t\right). \tag{5.14}$$

Lemma 5.2 implies that

$$\hat{P}_t\left(\{z^{(1)}, z^{(2)}\} \subset W_h \mid \bar{\Phi}_t^{(3)} = \bar{\Psi}_t\right) = (I_h)^2 \tag{5.15}$$

with

$$I_h = \frac{\int_{S^{d-1}} \mathrm{d}\nu \int_{\mathbb{R}^d} \mathrm{d}v'\, f_0(z_t)\, \mathbf{1}_{W_h}(z_t)\, \mathbf{1}_\Psi(z_t)}{\int_{S^{d-1}} \mathrm{d}\nu \int_{\mathbb{R}^d} \mathrm{d}v'\, f_0(z_t)\, \mathbf{1}_\Psi(z_t)}.$$

The estimation of I_h is straightforward: formula (5.7) implies that

$$\int \mathrm{d}f_0(u,v)\, \mathbf{1}_{V_h}(u,v) = a^{d-1}\, h^{2d}\, \left(f(u',v')\, [(v-v')\cdot\nu]_+ + o(1)\right) \tag{5.16}$$

as h tends to 0. Moreover, since $\Psi \in \mathcal{G}$ one obtains that there exists a constant C depending on Ψ such that

$$\int \mathrm{d}f_0(u,v)\, \mathbf{1}_\Psi(u,v) \geq 1 - \kappa_d\, C\, K_\infty\, a^{d-1} \geq \frac{1}{2}, \tag{5.17}$$

if a is sufficiently small. As a consequence of the bounds (5.16) and (5.17) the right hand side in estimate (5.15) tends to 0 as h tends to 0. Hence, we have established (5.14) and thus (5.13). Combining (5.12) and (5.13) one obtains eqn. (5.4) in the case that $\tau = t$.
Assume next that $\tau(\Phi) < t$ and define for each ν the set of colliding initial values with impact parameter ν:

$$W_h(\nu) = \left\{(u',v') \in U \times \mathbb{R}^d \;\middle|\; \begin{array}{l} \exists(\nu',t') \in S^{d-1}\times\mathbb{R} \text{ such that } u - u' + t'(v-v') + a\nu' = 0 \\ \text{and } (v-v')\cdot\nu' \geq 0 \text{ and } |t-t'| \leq \frac{h}{2} \text{ and } |\nu-\nu'| \leq \frac{h}{2} \end{array}\right\}.$$

We will show that

$$\lim_{h\to 0} h^{-2d}\hat{P}_t(\#(\omega \cap W_h(\nu)) > 0 \mid \Psi) = (1-\gamma)J \tag{5.18}$$

with

$$J = \frac{\int_{S^{d-1}} \mathrm{d}\nu \int_{\mathbb{R}^d} \mathrm{d}v'\, f_0(z_t)\, \mathbf{1}_\Psi(z_t)\, [(v-v')\cdot\nu]_+}{\int_{S^{d-1}} \mathrm{d}\nu \int_{\mathbb{R}^d} \mathrm{d}v'\, f_0(z_t)\, \mathbf{1}_\Psi(z_t)}.$$

Analogously to the case $\tau(\Phi) = t$ we obtain the upper and lower bound

$$0 \leq (1-\gamma)\, J - h^{-2d}\hat{P}_t(\#(\omega \cap W_h(\nu)) > 0 \mid \Psi) \leq o(1)$$

as $h \to 0$ with

$$J = \frac{1}{\int \mathrm{d}f_0(z)\, \mathbf{1}_\Psi(z)} \int_{\mathbb{R}^d} \mathrm{d}v'\, \mathbf{1}_\Psi(z)\, f_0(u + t(v-v') + a\nu, v')\, (v-v')\cdot\nu_+.$$

Thanks to (5.17) and the fact that J does not depend on h equation (5.18) holds. Thus we have established the claim also in the case $\tau < t$ and the proof is finished.
The assumption that $\Gamma = \emptyset$ does not involve a loss of generality. For generic Γ can be repeated line by line, except that we have to add the conditioning encoded by Γ to all expressions involving $\hat{P}$. □

5.1. **Convergence.** We now proceed to estimate the difference between the empiric distribution $\hat{P}_t^a$ and the idealized P_t. The key estimate which provides a quantitative link between Q and $\hat{Q}$ is provided by the following comparison principle.

Proposition 5.3. *Let $\Phi \in \mathcal{MT}$, $\Psi \in \mathcal{MT}^\alpha$ for some $\alpha > 0$ such that $(\Phi,\Psi) \in \mathcal{G}^{\alpha+1}$. Then the estimate*

$$1 - \frac{\hat{P}_t(\Phi \mid \Psi)}{P_t(\Phi)} \leq \rho(\#\Phi)\eta_t \tag{5.19}$$

holds for all $\Phi \in \mathcal{G}$*, with* $\rho(k) = (2k-1)(Ct\exp(Ct))^k$*,*

$$\eta_t(\Phi, \Psi) = \int f_0(u,v)\,(1 - \mathbf{1}_{\Phi,t}(u,v)\,\mathbf{1}_{\Psi,t}(u,v))(1+|v|)\,\mathrm{d}u\,\mathrm{d}v,$$
$$C(\Phi) = 2\max\left\{\sup\left\{\mu_t(u,v) \;:\; u \in U, |v| \le \max_{l\in m(\Phi)} |v_l|\right\}, \max\left\{|v(l) - v(l')| \;:\; l, l' \in m(\Phi)\right\}\right\},$$

where μ_t *is defined in eqn. (4.12).*

Note that only η_t is affected by the conditioning.

Proof. Observe that ρ is superadditive, i.e. $\rho(x+y) \ge \rho(x) + \rho(y)$. We use induction over α. The case $\alpha = 1$ will be treated below. If $\alpha > 1$, then we can split Φ into two trees $\Phi' \in \mathcal{MT}^1$ and $\Phi'' \in \mathcal{MT}^{\alpha-1}$. Clearly

$$\begin{aligned}\hat{P}_t(\Phi \mid \Psi) =& \hat{P}_t(\Phi' \mid \Phi'', \Psi)\hat{P}_t(\Phi'' \mid \Psi) \ge (1 - \rho(\#\Phi')\eta_t)(1 - \rho(\#\Phi'')\eta_t)P_t(\Phi' \mid \Phi'')P_t(\Phi'')\\ \ge& (1 - (\rho(\#\Phi') + \rho(\#\Phi''))\eta_t)P_t(\Phi) \ge (1 - \rho(\#\Phi)\eta_t)P_t(\Phi).\end{aligned}$$

The first inequality holds because of the induction assumption, the second one is due to the sign of the quadratic term and the third inequality is a consequence of the superadditivity of the function ρ. Thus, it suffices to consider the case $\alpha = 1$.
First, note that the definitions of L_t, $\hat{L}_t$ in (4.7) and (5.2) imply that for every probability measure P with marginal $\mu_t(u,v) = f_0(u - tv, v)$ the inequality

$$(1-\gamma)\hat{L}_t[P](\Phi) \ge L_t^a[P](\Phi)\begin{cases} 1 + 2\eta & \text{if } \tau(\Phi) \ne t,\\ 1 - \gamma & \text{if } \tau(\Phi) = t\end{cases} \tag{5.20}$$

holds, where $\tau > 0$ is the time of the final collision of Φ. If $t > \tau$ then the first term in (5.20) is relevant and one obtains

$$\begin{aligned}\frac{\partial}{\partial t}(\hat{P}_t - P_t^a)(\Phi) =& (1-\gamma)\hat{L}_t[\hat{P}_t](\Phi)\hat{P}_t(\Phi) - L_t^a[P_t^a](\Phi)P_t^a(\Phi)\\ \ge& (1+2\eta)L_t^a[\hat{P}_t](\Phi)\hat{P}_t(\Phi) - L_t^a[P_t^a](\Phi)P_t^a(\Phi)\\ =& (1+2\eta)L_t^a[\hat{P}_t](\Phi)\left(\hat{P}_t(\Phi) - P_t^a(\Phi)\right)\\ &+ (1+2\eta)(L_t^a[\hat{P}_t](\Phi) - L_t^a[P_t](\Phi))P_t(\Phi) + 2\eta L_t^a[P_t^a](\Phi)P_t^a(\Phi).\end{aligned} \tag{5.21}$$

For fixed Φ this is a one-dimensional ordinary differential inequality for $x(t) = (\hat{P}_t - P_t^a)(\Phi)$. Then we also obtain an integrated estimate. For one-dimensional ode, if $\dot{x}(t) \ge a(t)x(t) + b(t)$ and $x(\tau) = x_0$, i.e.

$$\begin{aligned}\dot{x}(t) =& a(t)x(t) + b(t) + c(t) \text{ with } c(t) \ge 0, x(\tau) = x_0\\ \dot{y}(t) =& a(t)y(t) + b(t) \text{ with } y(\tau) = x_0,\end{aligned}$$

then by the variation of constants formula we have

$$\begin{aligned}x(t) =& \exp\left(\int_\tau^t a(s)\,\mathrm{d}s\right)x_0 + \int_\tau^t \exp\left(\int_s^t a(\sigma)\,\mathrm{d}\sigma\right)(b(s) + c(s))\,\mathrm{d}s\\ \ge y(t) =& \exp\left(\int_\tau^t a(s)\,\mathrm{d}s\right)x_0 + \int_\tau^t \exp\left(\int_s^t a(\sigma)\,\mathrm{d}\sigma\right)b(s)\,\mathrm{d}s\end{aligned}$$

for $t \ge \tau$, independent of the signs of a and b. Then estimating

$$|\hat{\mu}_s - \mu_t|_{L^\infty(U\times\mathbb{R}^d)} \le 2\eta_s$$

and observing that $s \mapsto \eta_s$ is non-decreasing we obtain

$$\begin{aligned}
&\hat{P}_t(\Phi) - P_t(\Phi) \\
&\geq \exp\left((1+2\eta_t)\int_\tau^t L_s^a[\hat{P}_s]\,\mathrm{d}s\right)\left(\hat{P}_\tau(\Phi) - P_\tau^a(\Phi)\right) \\
&\quad + \int_\tau^t (1+2\eta_t)2\eta_t \exp\left((1+2\eta_t)\int_s^t L_{s'}^a[\hat{P}_{s'}](\Phi)\,\mathrm{d}s'\right) P_s^a(\Phi)\,\mathrm{d}s \\
&\quad + 2\eta_t \int_\tau^t \exp\left((1+2\eta_t)\int_s^t L_{s'}^a[\hat{P}_{s'}^a](\Phi)\,\mathrm{d}s'\right) L_s^a[P_s^a](\Phi)P_s^a(\Phi)\,\mathrm{d}s.
\end{aligned}$$

After observing that

$$\exp\left(\int_s^t L_{s'}^a[P_{s'}^a]\,\mathrm{d}s'\right) P_s^a(\Phi) = P_t^a(\Phi)$$

we obtain

$$\begin{aligned}
&\hat{P}_t(\Phi) - P_t(\Phi) \\
&\geq \exp\left((1+2\eta_t)\eta_t\int_\tau^t L_s^a[\hat{P}_s]\,\mathrm{d}s\right)\left(\hat{P}_\tau(\Phi) - P_\tau^a(\Phi)\right) \\
&\quad + \Big[\int_\tau^t (1+2\eta_t)2\eta_t \exp\left((1+2\eta_t)\int_s^t L_{s'}^a[\hat{P}_{s'}](\Phi)\,\mathrm{d}s'\right) \exp\left(-\int_s^t L_{s'}^a[P_{s'}^a](\Phi)\,\mathrm{d}s'\right)\mathrm{d}s \\
&\quad + 2\eta_t \int_\tau^t \exp\left((1+2\eta_t)\int_s^t L_{s'}^a[\hat{P}_{s'}]\,\mathrm{d}s'\right) L_s^a[P_s^a](\Phi) \exp\left(-\int_s^t L_{s'}^a[P_{s'}^a](\Phi)\,\mathrm{d}s'\right)\mathrm{d}s\Big] P_t^a(\Phi).
\end{aligned} \tag{5.22}$$

We use induction over $k = \#\Phi$. First assume that $\#\Phi = 1$. In this case $\tau = 0$ and $P_\tau(\Phi) = \hat{P}_\tau(\Phi)$, which together with (5.22) implies that

$$\begin{aligned}
&\hat{P}_t(\Phi) - P_t^a(\Phi) \\
&\geq \Big[\int_\tau^t (1+2\eta_t)2\eta_t \exp\left((1+2\eta)\int_s^t L_{s'}^a[\hat{P}_{s'}]\,\mathrm{d}s'\right) \exp\left(-\int_s^t L_{s'}^a[P_{s'}^a](\Phi)\,\mathrm{d}s'\right)\mathrm{d}s \\
&\quad + 2\eta \int_\tau^t \exp\left((1+2\eta_t)\int_s^t L_{s'}^a[\hat{P}_{s'}]\,\mathrm{d}s'\right) L_s^a[P_s^a](\Phi) \exp\left(-\int_s^t L_{s'}^a[P_{s'}^a](\Phi)\,\mathrm{d}s'\right)\mathrm{d}s\Big] P_t^a \\
&\geq -Ct\exp(Ct)\eta_t P_t^a(\Phi).
\end{aligned}$$

Assume next that the estimate has been established for all trees with at most k nodes and let $\#\Phi = k+1$. Define $k_1 = \#\bar{\Phi}_\tau$ and $k_2 = \#\Phi'_\tau$ with $k_1 + k_2 = k+1$ and $\max\{k_1, k_2\} \leq k$. Thus the induction assumption implies hat

$$\begin{aligned}
&\hat{P}_\tau(\bar{\Phi}_\tau) \geq (1-\rho(k_1))P_\tau^a(\bar{\Phi}_\tau), \\
&\hat{P}_\tau(\Phi'_\tau \mid \bar{\Phi}_\tau) \geq (1-\rho(k_2))P_\tau^a(\Phi'_\tau).
\end{aligned}$$

Using equation (5.22) we find that

$$\begin{aligned}
&\hat{P}_t(\Phi) - P_t^a(\Phi) \\
&\geq \exp\left((1+2\eta)\int_\tau^t L_s^a[\hat{P}_s]\,\mathrm{d}s\right)\Big(\underbrace{\hat{P}_\tau(\bar{\Phi}_\tau)\hat{P}_\tau(\Phi'_\tau \mid \bar{\Phi}_\tau)}_{\geq P_\tau(\bar{\Phi}_\tau)P_\tau(\Phi'_\tau)(1-(\rho(k_1)+\rho(k_2)))} - P_\tau(\bar{\Phi}_\tau)P_\tau(\Phi'_\tau)\Big) \\
&\quad + \Big[\int_\tau^t (1+2\eta)2\eta \exp\left((1+2\eta)\int_s^t L_{s'}^a[\hat{P}_{s'}](\Phi)\,\mathrm{d}s'\right) \exp\left(-\int_s^t L_{s'}^a[P_{s'}^a](\Phi)\,\mathrm{d}s'\right)\mathrm{d}s \\
&\quad + 2\eta \int_\tau^t \exp\left((1+2\eta)\int_s^t L_{s'}^a[\hat{P}_{s'}^a]\,\mathrm{d}s'\right) L_s^a[P_s^a](\Phi) \exp\left(-\int_s^t L_{s'}^a[P_{s'}^a](\Phi)\,\mathrm{d}s'\right)\mathrm{d}s\Big] P_t^a \\
&\geq -(\rho(k_1)+\rho(k_2)) + Ct\exp(Ct)\eta)P_t(\Phi).
\end{aligned}$$

Since

$$(\rho(k_1)+\rho(k_2)+1)Ct\exp(Ct)\le Ct\exp(Ct)(2k_1+2k_2-2+1)(Ct\exp(Ct))^{\max\{k_1,k_2\}}\le\rho(k+1),$$

inequality (5.19) has been established. ☐

The estimate (5.19) immediately implies that $\hat{P}_t$ converges in the total variation sense to P_t as a tends to 0.

Proposition 5.4. *For each $\alpha\ge 1$ the total variation distance between P^α and $\hat{P}^\alpha$ satisfies*

$$\lim_{a\to 0}\|P_t^{a,\alpha}-\hat{P}_t^\alpha\|_{L^1(\mathcal{MT})}=0$$

Proof. We assume that $\alpha=1$. Lemma 4.12 implies that for each $\varepsilon>0$ there exists $a>0$ so small that

$$P_t^a(\mathcal{MT}\setminus\mathcal{G}(a))\le\varepsilon.$$

We use that

$$\|P_t^a-\hat{P}_t\|=2\sup_\Omega(P_t^a(\Omega)-\hat{P}_t(\Omega)).$$

For each Ω one has

$$\begin{aligned}P_t^a(\Omega)-\hat{P}_t(\Omega)=&P_t^a(\Omega\cap\mathcal{G}(a))-\hat{P}_t(\Omega\cap\mathcal{G}(a))+P_t^a(\Omega\setminus\mathcal{G}(a))-\hat{P}_t(\Omega\setminus\mathcal{G}(a))\\ \le&\rho\zeta(a)P_t^a(\mathcal{MT})+P_t^a(\mathcal{MT}\setminus\mathcal{G}(a))\le\rho\zeta+P_t^a(\mathcal{MT}\setminus\mathcal{G}(a)),\end{aligned}\tag{5.23}$$

where $\zeta(a,a_0):=\sup_{\Phi,\Psi\in\mathcal{G}(a_0)}\eta_t(\Phi,\Psi)$ with ρ and $\eta_t(\Phi,\Psi)$ defined in Proposition 5.3. Uniformly for $\Phi,\Psi\in\mathcal{G}(a_0)$ a particle in Φ,Ψ can at most cover a cylinder of volume $V(a_0)t\kappa_d a^{d-1}+\kappa_{d+1}a^d$, where the second term describes the initial overlap. There at most $2M(a_0)$ particles in $\Phi\cup\Psi$ and the u-marginal of f_0 in bounded in $L^\infty(U)$ by K_∞, hence

$$\zeta(a,a_0)\le 2M(a_0)V(a_0)t\kappa_d a^{d-1}K_\infty\tag{5.24}$$

The function ρ is uniformly bounded on $\mathcal{G}(a_0)$ by definition. Then $\lim_{a\to0}\zeta(a,a_0)=0$ for any fixed a_0 and $\lim_{a\to0}P_t^a(\mathcal{MT}\setminus\mathcal{G}(a))=0$ (Lemma 4.12), this implies the claim. ☐

6. Effective dynamics

Proof of Theorem 2.1. We first show that the distribution of a single tagged particle satisfies the gainless Boltzmann equation (2.5). Let $A\subset U\times\mathbb{R}^d$ and define $\Omega_t(A)\subset\mathcal{MT}$ by

$$\Omega_t(A)=\{\Phi\;:\;\beta(\Phi)=1\text{ and }(u+tv,v)\in A\}.\tag{6.1}$$

According to Proposition 4.9 every weak solution f_t of equation (4.29) is a mild solution and thereby unique. Proposition 4.8 implies that

$$\int_A\mathrm{d}f_t=P_t(\Omega_t),$$

and thus

$$\left|\lim_{a\to0}\hat{P}_t(\Omega_t)-\int_A\mathrm{d}f(u,v,t)\right|\overset{\text{Prop. 4.8}}{=}\lim_{a\to0}\left|\hat{P}_t(\Omega)-P_t(\Omega)\right|\overset{\text{Prop. 5.4}}{=}0.$$

The convergence is uniform in A by (5.23).
Next we consider the random variables for $i\in\{1,\dots,n\}$

$$\chi_i(t)=\begin{cases}1&\text{if }(u_i(t),v_i(t))\in A\text{ and }\beta_i^{(a)}(t)=1\\0&\text{else.}\end{cases}$$

Then the first part of the proof yields that $\lim_{a\to 0}\langle\chi_i(t)\rangle = \int_A \mathrm{d}f_t(u,v)$ for each i. Now we define the random variable $s_n = \frac{1}{n}\sum_{i=1}^n \chi_i(t)$. The claim (2.9) follows if the variance $V_n = \langle(s_n - \langle s_n\rangle)^2\rangle$ converges to 0 as n tends to infinity. Thanks to the permutation invariance we obtain that

$$\begin{aligned} V_n \leq& \frac{1}{n}(\chi_1(t) - \langle\chi_1(t)\rangle)^2 + \frac{n-1}{n}\langle(\chi_1(t) - \langle\chi_1(t)\rangle)(\chi_2(t) - \langle\chi_2(t)\rangle)\rangle \\ \leq& \frac{1}{n} + |\langle(\chi_1(t) - \langle\chi_1(t)\rangle)(\chi_2(t) - \langle\chi_2(t)\rangle)\rangle|\,. \end{aligned}$$

If we apply Proposition 5.4 again with $\alpha = 2$ and

$$\Omega_t(A,B) = \left\{ \Phi \in \mathcal{MT}^2 \;\middle|\; \beta_1(\Phi) = \beta_2(\Phi) = 1, (u_1 + tv_1, v_1) \in A \text{ and } (u_2 + tv_2, v_2) \in B \right\}$$

we obtain that

$$\lim_{a\to 0}\langle\chi_1(t)\chi_2(t)\rangle = \lim_{a\to 0}\hat{P}_t^2(\Omega(A,A)) = P_t(\Omega(A))^2.$$

This implies that

$$\lim_{a\to 0}\langle(\chi_1(t) - \langle\chi_1(t)\rangle)(\chi_2(t) - \langle\chi_2(t)\rangle)\rangle = 0 \tag{6.2}$$

uniformly in A which completes the proof of (2.9). Equation (6.2) is the main reason to consider trees with several roots. In particular this gives $V_n \leq b(n)$ for some decaying function $b : \mathbb{N} \to \mathbb{R}$ uniformly in the test set A again by (5.23).
Finally we show (2.10). We recall a well-known principle in probability theory. Let $x_N \in \mathbb{R}$ be a sequence of independent random numbers such that $\mathbb{E}(x_N) = 0$ and let V_N be the variance of x_N. If $\sum_{N=1}^\infty V_N < \infty$, then almost surely $\lim_{N\to\infty} x_N = 0$.
Indeed, for every $\varepsilon, N_0 > 0$ Tchebychev's inequality yields the estimate

$$\mathrm{Prob}\left(\sup_{N\geq N_0} |x_N| \leq \varepsilon\right) \geq \prod_{N=N_0}^{\infty}\left(1 - \tfrac{V_N}{\varepsilon^2}\right) \geq 1 - \frac{1}{\varepsilon^2}\sum_{N=N_0}^{\infty} V_N.$$

Hence $\lim_{a\to 0}\mathrm{Prob}(\sup_{N\geq N_0}|x_N| \leq \varepsilon) = 1$, i.e. for each realization and each $\varepsilon > 0$ there exists almost surely a number $N_0 > 0$ such that $\sup_{N\geq N_0}|x_N| \leq \varepsilon$.
Let s_n be the sum and V_n be the variance of s_n as above. Since $\lim_{a\to 0} V_n = 0$ uniformly in A there exists a subsequence V_{n_k} such that $\sum_{k=1}^\infty V_{n_k} < \infty$ for all A. We apply now the previous consideration to the sequence $x_k = s_{n_k}$ such that

$$\int_A \frac{1}{n_k}\sum_{i=1}^{n_k}\beta_i^{(a_k)}(t)\,\delta(\,\cdot\, - (u_i(t), v_i(t)))\,\mathrm{d}u\,\mathrm{d}v \to \int_A \mathrm{d}f_t$$

as k tends to infinity and thus we obtain the desired weak-$*$ convergence (2.10). □

7. Spatial concentrations

We discuss variants and limitations of the presented theory. First we consider the regularity assumptions on the initial density f_0.
We require $\int_{\mathbb{R}^d} f_0(\,\cdot\,,v)\,\mathrm{d}v \in L^\infty(U)$ in (1.5), this implies that in eqn. (5.24) the expected number of particles overlapping with any given particle converges to 0 with a. The result holds also with less restrictive assumptions on the initial distribution.

Proposition 7.1. *Let $\int_{\mathbb{R}^d} f_0(\,\cdot\,,v)\,\mathrm{d}v \in L^d(U)$, then the expected number of overlapping particles at a given point u_0 converges to zero for $a \to 0$.*

Proof. The expected number of particles in ball B_a of radius a around a $u_0 \in U$ is given by $p(u_0) = n\int_{B_a(u_0)}\int_{\mathbb{R}^d}\mathrm{d}f_0(u,v)$, which by the scaling and (2.8) and can be estimated using the

Hölder inequality

$$p(u_0) \leq n \left(\int_{B_a} 1 \, \mathrm{d}u \right)^{(d-1)/d} \left(\int_{B_a} \left(\int_{\mathbb{R}^d} \mathrm{d}v f_0(u,v) \right)^d \mathrm{d}u \right)^{1/d}$$

$$= a^{1-d} \left(\kappa_{d+1} \, a^d \right)^{(d-1)/d} \left(\int_{B_a} \left(\int_{\mathbb{R}^d} \mathrm{d}v f_0(u,v) \right)^d \mathrm{d}u \right)^{1/d} \to 0$$

for $a \to 0$ as $\int_{\mathbb{R}^d} \mathrm{d}v \, f_0(\,\cdot\,, v) \in L^d(U)$. □

Whereas for $\int_{\mathbb{R}^d} \mathrm{d}v f_0(u,v) = |u|^p$ near $u = 0$ with $-d < p < -1$, we still have $\int_{\mathbb{R}^d} \mathrm{d}v f_0(u,v) \in L^1(U)$ but the expected number of particles in a ball of radius a around 0 tends to infinity, but this effect will not be statistically relevant, as the growth is sublinear in n.
We now modify this example such that the expected number of nodes in the empiric trees tends to infinity for a tending to zero, even for short times when $f_0 \in L^1_{1+|v|}(U \times \mathbb{R}^d)$. Note that the idealized theory leads to the integral equation (4.30), which is well-defined for initial data in $L^1_{1+|v|}(U \times \mathbb{R}^d)$ and can be easily interpreted for measures. While the first example does not show non-validity due to singularities, it highlights difficulties in a proof for more general initial distributions, as tightness (4.27) was crucial to restrict the error estimates to trees of finite height.

Proposition 7.2. *There exists an initial distribution $f_0 \in L^1_{1+|v|}(U \times \mathbb{R}^d)$ such that the expected number of overlaps $\int_{U \times \mathbb{R}^d} p(u) f_0(u,v)$ is unbounded as $a \to 0$.*

Proof. Let $(u_i)_{i \in \mathbb{N}^d}$ be an ordering of all diadic fractions on the torus U such that for every pair i, j with

$$u_i = \left(\frac{i_1}{2^{k_1}}, \ldots, \frac{i_d}{2^{k_d}} \right) \qquad \text{with } \gcd(i_1, 2^{k_1}) = \ldots = \gcd(i_d, 2^{k_d}) = 1,$$

$$u_j = \left(\frac{j_1}{2^{l_1}}, \ldots, \frac{j_d}{2^{l_d}} \right) \qquad \text{with } \gcd(j_1, 2^{l_1}) = \ldots = \gcd(j_d, 2^{l_d}) = 1,$$

and $\max\{k_1, \ldots, k_d\} > \max\{l_1, \ldots, l_d\}$ then $i > j$. We consider

$$f_0(u,v) = c \sum_{j=1}^{\infty} c_j |u - u_j|^p \bar{f}(v) \tag{7.1}$$

with $-d < p < -1$ to be chosen later. The density $\bar{f} \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ is non-negative and normalized ($\int_{\mathbb{R}^d} \mathrm{d}\bar{f} = 1$) and the constant c is chosen such that $\int_{U \times \mathbb{R}^d} \mathrm{d}f_0 = 1$. We will consider in particular the cases $d \geq 3$. Proposition 4.8 yields existence of a solution to (4.30), which is at least $L^1_{1+|v|}(U \times \mathbb{R}^d)$, when $f_0 \in L^1_{1+|v|}(U \times \mathbb{R}^d)$. For f_0 as in (7.1) there exists a $p \in (-d, -1)$ such that empiric expected number of overlapping particles is unbounded.
The expected number of particles overlapping the first particle is $\int_{U \times \mathbb{R}^d} \mathrm{d}f_0(u,v) \, p(u)$. If $u \in B_{a/2}(u_j)$ then for some constant C independent of j and a:

$$p(u) = n \int_{B_a(u) \times \mathbb{R}^d} \mathrm{d}f_0(u', v) \geq nc \int_{B_a(u)} c_j |u' - u_j|^p \, \mathrm{d}u' \geq nC \int_{B_{a/2}(u_j)} c_j |u' - u_j|^p \, \mathrm{d}u'$$

$$\geq a^{1-d} C \, c_j \int_0^{a/2} r^p \, r^{d-1} \, \mathrm{d}r = C \, c_j \, a^{1+p}.$$

We choose $J(a) \in \mathbb{N}$ such that the balls $B_{a/2}(u_j)$ are disjoint for $j = 1, \ldots, J$. The expected number of overlaps can be bounded from below by

$$\int_{U\times\mathbb{R}^d} \mathrm{d}f_0(u,v)\,p(u) \geq \sum_{j=1}^{J} \int_{B_{a/2}(u_j)\times\mathbb{R}^d} \mathrm{d}f_0(u,v)\,p(u)$$
$$\geq \sum_{j=1}^{J} \left(\frac{a}{2}\right)^d c_j \left(\frac{a}{2}\right)^p C\, c_j\, a^{1+p} = Ca^{d+2p+1} \sum_{j=1}^{J} c_j^2,$$

where C is a constant that does not depend on a. For $-d < p < -\frac{d+1}{2}$ this is unbounded as $a \to 0$. □

We give now an example of non-validity if we allow concentrations in space and velocity space. Note that in the spatially homogeneous case we could prove validity unless there was concentration on lines. With $\mathcal{H}^2$ denoting 2-dimensional Hausdorff measure, we consider initial data concentrated on a sphere $S = \{u \in U \ : \ |u - u_0| = r\}$ where $r \in (0, \frac{1}{2})$ and $u_0 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$.

$$f_0(u,v) = c\mathcal{H}^2|_S \times \delta_0(v - (u - u_0)), \tag{7.2}$$

where c is a normalization constant.

Proposition 7.3. *Let f_0 be given by (7.2) and let the solution of the gainless Boltzmann equation (2.5) given by (4.30), then the convergence in probability (2.9) does not hold.*

Proof. We first observe for the idealized prediction that $f_t(u,v) \leq f_0(u - tv, v)$ by (4.30). Now suppose $f(u - tv, v) \neq 0$ for some $t \in (0, 1/4)$ then $f_0(u - (t - s)v, v') = 0$ for all $v' \in \mathbb{R}^d$ and all $0 < s < t$. Then $f_s(\tilde{u}, v') = 0$ for all $\tilde{u} = u - (t - s)v$ and all v' by (4.30). Hence the integral in the argument of the exponential in (4.30) vanishes and $f_t(u,v) = f_0(u - tv, v)$. This is equivalent to pure transport until $t = \frac{1}{4}$. The measure is concentrated on a decreasing sphere in u and the two-dimensional Hausdorff measure is scaled by $\left(\frac{r(t)}{r(0)}\right)^2$.
On the other hand the hardsphere flow is well-defined with f_0 being a general measure. The n particles are distributed randomly with uniform distribution on a shrinking two-dimensional sphere of radius $\frac{1}{4} - t$ and surface area $4\pi(\frac{1}{4} - t)^2$.
If an iid-distribution is used, then a macroscopic portion of the particles is instantaneously removed almost surely thanks to the definition of the scattering state $\beta_i^{(a)}(0)$ in eqn. (2.4). The following calculation shows that even in the case of more general distributions a macroscopic proportion of the particles undergoes a collision before time $t = \frac{1}{4}$.
The surface covered by n balls of diameter a is $na^2\pi/4 = \pi/4$ by (2.8), i.e. at most the fraction $\frac{4\pi(1/4-t)^2}{\pi/4} = 16(1/4 - t)^2 = 1 - 8t + t^2$ of particles has not collided by time $t \in (0, 1/4)$. Thus for any empiric distribution $\frac{1}{n}\#\left\{i \;\middle|\; (u_i(t), v_i(t)) \in U \times \mathbb{R}^d,\, \beta_i^{(a)}(t) = 1\right\} \leq 1/4$ for $1/8 < t < 1/4$. □

References

[Ag84] V. Agoshkov. Spaces of functions with differential-difference characteristics and the smoothness of solutions of the transport equation, *Dokl. Akad. Nauk SSSR* **276** (1984), no. 6, 1289–1293.

[Ba77] J. Ball. Strongly continuous semigroups, weak solutions, and the variation of constants formula. *Proc. Amer. Math. Soc.* **63** (1977), 370–373.

[BGL93] C. Bardos, F. Golse, C. Levermore. Fluid dynamic limits of kinetic equations. II. Convergence proofs for the Boltzmann equation. *Comm. Pure Appl. Math.* **46** (1993), 667–753.

[BBS83] C. Boldrighini, L.A. Bunimovich, Y.G. Sinai. On the Boltzmann equation for the Lorentz gas. *J. Stat. Phys.* **32** (1983), 477-501.

[CIP94] C. Cercignani, R. Illner, M. Pulvirenti. *The Mathematical Theory of Dilute Gases.* Applied Mathematical Sciences, Vol **106**, Springer Verlag, New York (1994).

[CDPTW03] F. Coppex, M. Droz, J. Piasecki, E. Trizac, P. Wittwer. Some exact results for Boltzmann's annihilation dynamics. *Phys. Rev. E* **79** (2003), 21103.

[DM78] C. Dellacherie, P. Meyer. *Probabilities and potential.* North-Holland Mathematics Studies, **29**. North-Holland Publishing Co., Amsterdam-New York, (1978).

[DP01] R.DeVore, G. Petrova. The averaging lemma, *J. Amer. Math. Soc.* **14** (2001), 279–296.
[DL89a] R. DiPerna, P.L. Lions. On the Cauchy Problem for Boltzmann Equations: Global Existence and Weak Stability, *Ann. Math.* **130** (1989) 321–366.
[DL89b] R. DiPerna, P.L. Lions. Ordinary differential equations, transport theory and Sobolev spacesl, *Invent. Math.* **98** (1989), 511–547.
[DLM91] R. DiPerna, P.L. Lions, Y. Meyer. L^p regularity of velocity averages, *Ann. Inst. H. Poincaré, Analyse non-linéaire* **8** (1991), 271–287.
[DFPR95] M. Droz, L. Frachebourg, J. Piasecki, P.-A. Rey. Ballistic annihilation kinetics for a multivelocity one-dimensional ideal gas. *Phys. Rev. E* **51** (1995), 5541–5548.
[DP91] E. Dawson, E. Perkins. *Historical processes.* American Mathematical Society, 1991.
[EF85] Y. Elskens, H. Frisch. Annihilation kinetics in the one-dimensional ideal gas. *Phys. Rev. A* **31** (1985), 3812 – 3816.
[Ge90] P. Gérard. Moyennisation et régularité deux-microlocale. *Ann. Sci.Ecole Norm. Sup.*(4) **23** (1990), 89–121.
[GG92] P. Gérard F. Golse. Averaging regularity results for PDEs under transversality assumptions. *Comm. Pure Appl. Math.* **45** (1992), 1–26.
[GLPS99] F. Golse, P.-L. Lions, B. Perthame, R. Sentis. Regularity of the moments of the solution of a transport equation. *J. Funct. Anal.* **76** (1988), 110–125.
[GS04] F. Golse, L. Saint-Raymond. The Navier-Stokes limit of Boltzmann equation: convergence proof. *Invent.Math.* **155** (2004), 81–161.
[IP89] R. Illner, M. Pulvirenti. Global validity of the Boltzmann equation for two- and three-dimensional gas in vacuum. Erratum and improved result. *Comm. Math. Phys* **121** (1989), 143–146.
[KS88] J. Krug, H. Spohn. Universality classes for deterministic surface growth. *Phys. Rev. A* **38** (1988), 4271–4283.
[Lan75] O. Lanford. *Time evolution of large classical systems.* ed. by J. Moser pp 1-113. Lecture notes in physics, Vol **38**, Springer Verlag (1975).
[LN80] R. Lang, X. Nguyen. Smoluchowski's theory of coagulation holds rigorously in the Boltzmann-Grad limit. *Z. Wahrs. Verw. Geb.* **54** (1980), 227–280.
[LM01a] P.-L. Lions, N. Masmoudi. From the Boltzmann equations to the equations of incompressible fluid mechanics. I. *Arch. Ration. Mech. Anal.* **158** (2001), 173–193.
[LM01b] P.-L. Lions, N. Masmoudi. From the Boltzmann equations to the equations of incompressible fluid mechanics. II. *Arch. Ration. Mech. Anal.* **158** (2001), 195–211.
[MT08] K. Matthies, F. Theil. Validity and non-validity of propagation of chaos in *Analysis and Stochastics of Growth Processes and Interface Models* edited by P. Mörters, R. Moser, M. Penrose, H. Schwetlick, J. Zimmer, OUP, 2008, 101-119.
[MT10] K. Matthies, F. Theil. Validity and failure of the Boltzmann approximation of kinetic annihilation, *J. Nonl. Sc.* **20** (2010), 1-46.
[Per04] B. Perthame. Mathematical tools for kinetic equations. *Bull. Amer. Math. Soc. (N.S.)* **41** (2004), 205–244.
[Paz83] A. Pazy. *Semigroups of linear operators and applications to partial differential equations.* Applied Mathematical Sciences, Springer-Verlag, New York (1983).
[Pia95] J. Piasecki. Ballistic annihilation in a one-dimensional fluid. *Phys. Rev. E* **51** (1995), 5535–5540.
[PTD02] J. Piasecki, E. Trizac, M. Droz. Dynamics of ballistic annihilation. *Phys. Rev. E* **65** (2002), 66111.
[Sha88] M. Sharpe. *General Theory of Markov Processes.* Academic Press (1988).
[Spo78] H. Spohn. The Lorentz process converges to a random flight process. *Comm. Math. Phys.* **60** (1978), 277-290.
[Spo91] H. Spohn. *Large scale dynamics of interacting particles.* Texts and Monographs in Physics, Springer Verlag (1991).

Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom
E-mail address: `k.matthies@maths.bath.ac.uk`

Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom
E-mail address: `theil@maths.warwick.ac.uk`